TRANSACTIONS
OF THE
AMERICAN INSTITUTE
OF
ELECTRICAL ENGINEERS

VOLUME 73
1954

PART I. COMMUNICATION AND ELECTRONICS

PUBLISHED BY THE
AMERICAN INSTITUTE OF ELECTRICAL ENGINEERS
33 WEST 39TH STREET
NEW YORK 18, N. Y., U. S. A.

Copyright 1954 by the American Institute of Electrical Engineers
Printed in U. S. A. by The Mack Printing Company, Easton, Pa.

Preface

The AIEE *Transactions* for 1954 (volume 73) is published in three parts: Part I. Communication and Electronics; Part II. Applications and Industry; and Part III. Power Apparatus and Systems. (Part III is divided into two parts: III-A, consisting of pages 1–832, and III-B, consisting of pages 833–1780. The Index for both parts appears at the end of III-B.) The papers in each of the three parts are classified according to subject matter as follows:

Part I. Communication and Electronics

- Communication Switching Systems
- Radio Communications Systems
- Special Communications Applications
- Telegraph Systems
- Television and Aural Broadcasting Systems
- Wire Communications Systems
- Basic Sciences
- Computing Devices
- Electrical Techniques in Medicine and Biology
- Electronic Power Converters
- Electronics
- Instruments and Measurements
- Magnetic Amplifiers
- Metallic Rectifiers
- Nucleonics

Part II. Applications and Industry

- Air Transportation
- Domestic and Commercial Applications
- Land Transportation
- Marine Transportation
- Production and Application of Light
- Chemical, Electrochemical and Electrothermal Applications
- Electric Heating
- Electric Welding
- Feedback Control Systems
- General Industry Applications
- Industrial Control
- Industrial Power Systems
- Mining and Metal Industry

Part III. Power Apparatus and Systems

- Carrier Current
- Insulated Conductors
- Power Generation
- Protective Devices
- Relays
- Rotating Machinery
- Substations
- Switchgear
- System Engineering
- Transformers
- Transmission and Distribution

- Education
- Safety
- Research

Each part has been indexed separately in the back of that particular part. The three parts are not cross-referenced; hence the user should determine first whether the subject matter of the paper desired is in the field of communication and electronics, applications and industry, or power apparatus and systems. Papers are listed in the subject index under several key words in the titles. The original numbers assigned to the papers are given in the author index. Volume 73 contains the technical papers and related discussions presented at these meetings:

1. Winter General Meeting, New York, N. Y., January 18–22, 1954.
2. North Eastern District Meeting, Schenectady, N. Y., May 5–7, 1954.
3. Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954.
4. Middle Eastern District Meeting, Reading, Pa., October 5–7, 1954.
5. Fall General Meeting, Chicago, Ill., October 11–15, 1954.

Statements and opinions given in the papers and discussions published in *Transactions* are the expressions of the contributors for which the American Institute of Electrical Engineers assumes no responsibility.

Communication and Electronics Index for 1954

1. Technical Subject Index

Acceleration Plane Method for Analysis of a Circuit with Nonlinear Inductance and Nonlinear Capacitance. Ku........619–26
Accurate Tachometry Methods with Electronic Counters. Shulman........452–5; disc. 456
Advances in the Economy of Routing Calls in Nationwide Toll Dialing, Some Recent. Molnar..670–9
(AIEE-IRE) Proposed Methods of Testing Transistors. (Committee Report) (January 1955 C & E only)........725–40
Air, The Measurement of Electrostatic Potential Attributable to Net Ion Space Charge in. Carroll, Hammond........712–15; disc. 715
Airborne Servo-Control Systems, An Operational-Type Magnetic Amplifier for. Hubbard........425–30; disc. 430
Algebra, A Rectifier. Schaefer........679–82
Alloy Junction Diode as a Reference Standard in Regulated Metallic Rectifier Circuits, The Suitability of the Silicon, Smith........645–51; disc. 651
Alteration of the Dynamic Response of Magnetic Amplifiers by Feedback. Decker........658–63; disc. 663
A-C Galvanometer, Basic Theory and Experimental Verification of the. Higgins, Kneen........235–41; disc. 241
Alternating Current, Eddy-Current Losses in a Semi-Infinite Solid due to a Nearby. Poritsky, Jerrard........97–106
Ammeters for Use at Very High Frequencies, Thermocouple-Type. McAninch........241–5; disc. 245
Amplifier—A Low-Level D-C Magnetic Amplifier, Pulse Relaxation. Morgan, McFerran.....245–9
Amplifier—A Pulse Magnetic Amplifier, The Biased Rectifier. Morgan........584–90
Amplifiers, An Ultrahigh-Frequency Transmitter Employing Klystron Power. Sayer........114–18
Amplifiers for Community Television Systems, Line. Simons, Kirk, Arbeiter........214–19; disc. 219
AN/FGC-29 Circuit Design Aspects, Telegraph Terminal. Boughtwood........531–6
AN/FGC-29 Equipment Features, Telegraph Terminal. Cusack........536–40
AN/TRC-24 Radio Set, Equipment and Mechanical Features of the. Cruser........544–7
Analogue Computer for Automatic Determination of System Swing Curves, An. Shen, Lisser........475–83; disc. 483
Analogue Computer, Rectifier Arc-Back Study on the. Dillard, Baldwin........198–207; disc. 207
Analogue Computer Technique Using Magnetic Amplifiers, An. Davis, Swift........635–9; disc. 639
Analogue Computers, Some Transistor Building Blocks for. Hellerman........410–13
Analogue Shaft-Position Transducers, Networks for Digital-to-. O'Neil........456–66
Analysis and Synthesis of Sampled-Data Control Systems. Jury........332–46
Analysis for Nonlinear Characteristics, Harmonic. Lewis........693–9; disc. 699
Analysis of a Circuit with Nonlinear Inductance and Nonlinear Capacitance, Acceleration Plane Method for. Ku........619–26
Analysis of a Single-Core Magnetic Amplifier with Real Rectifier and Core Functions. Frank, Rabotnick, Walker........182–8; disc. 188
Analysis of Linear Time-Varying Circuits by the Brillouin-Wentzel-Kramers Method. Pipes......93–6
Analysis of Magnetic Amplifiers by the Use of Difference Equations. Johannessen........700–11; disc. 711
Analysis of Nonlinear Coupled Circuits. Ku........626–30; disc. 630
Annular Wave Guide Rotary Joint with Wave Guide Feed, An. Breetz........62–6
Antenna, Torque Requirements of a Radar. Mark........15–18
Antennas for Ultrahigh-Frequency Television Broadcasting, High-Gain Side-Firing Helical. Smith........135–8
Application and Transmission Features of a New 12-Channel Open-Wire Carrier System. Appert, Caruthers, Chaskin........18–27
Application of Network Theory to the Analysis of Rotating Machinery. Koenig:
  Part I. Synchronous and Asynchronous Machines........162–9
  Part II. Commutating Machines........169–74
  Discussions........174
Application of Toll Dialing in Pennsylvania. Charny........400–05
Applications of Semiconductor Devices in the Feedback Loop of Regulated Metallic Rectifiers, Some. Hamilton........640–5
Approach to a Video SHF Relay System, One. Coe, McClatchie........323–9
Arc-Back Study on the Analogue Computer, Rectifier. Dillard, Baldwin........198–207; disc. 207
Arc with Various Electrode Materials, Characteristics of the High Current Argon. Dzimianski, Jones........665–9; disc. 669
Argon Arc with Various Electrode Materials, Characteristics of the High-Current. Dzimianski, Jones........665–9; disc. 669
Armed Services, High Speed Teletypewriter Equipment for the. Schultheiss........88–93
Automatic Determination of System Swing Curves, An Analogue Computer for. Shen, Lisser........475–83; disc. 483
Automatic Oscillograph and Application on the Bonneville Power System, A New 12-Element. Hathaway, Davis, Curtin........307–12
Automatic Percussion Welding. Quinlan........561–5
Automatic Phase Control, Closed-Loop. Ordung, Gibson, Shinn........375–81
Automatic Teletypewriter Switching Center for Military Use, A Fully. Johnston, Stiles........27–37
Automatic Voltage Control System for Electrical Precipitators, An. Hall........124–7
Automatization in the Swedish Telephone System, Rural. Bjurel........552–61

B

(Basic Sciences) A Comparison of Metals and Ferrites for High-Speed Pulse Operation. Brown, Buck, Menyuk........631–4; disc. 634
(Basic Sciences) A Rectifier Algebra. Schaefer..679–82
(Basic Sciences) A Symbolic Method for Synthesis of 2-Terminal Switching Circuits. Zeheb, Caywood........690–3
(Basic Sciences) Acceleration Plane Method for Analysis of a Circuit with Nonlinear Inductance and Nonlinear Capacitance. Ku........619–26
(Basic Sciences) Analysis and Synthesis of Sampled-Data Control Systems. Jury........332–46
(Basic Sciences) Analysis of Linear Time-Varying Circuits by the Brillouin-Wentzel-Kramers Method. Pipes........93–6
(Basic Sciences) Analysis of Nonlinear Coupled Circuits. Ku........626–30; disc. 630
(Basic Sciences) Application of Network Theory to the Analysis of Rotating Machinery. Koenig:
  Part I. Synchronous and Asynchronous Machines........162–9
  Part II. Commutating Machines........169–74
  Discussions........174
(Basic Sciences) Calculation of Life Characteristics of Insulation. Whitman, Doigan....193–7; disc. 197
(Basic Sciences) Characteristics of the High-Current Argon Arc with Various Electrode Materials. Dzimianski, Jones........665–9; disc. 669
(Basic Sciences) Eddy-Current Losses in a Semi-Infinite Solid due to a Nearby Alternating Current. Poritsky, Jerrard........97–106
(Basic Sciences) Eddy-Current Phenomena in Ferromagnetic Materials. McConnell....226–33; disc. 233
(Basic Sciences) Effect of Electric Discharges on the Breakdown of Solid Insulation. Dakin, Philofsky, Divens........155–61; disc. 161
(Basic Sciences) Equations for Determining Current Distribution Among the Conductors of Busses Comprised of Double Channel Conductors. Siegel, Higgins........489–92
(Basic Sciences) Flow of Energy in Synchronous Machines. Hawthorne........1–10
(Basic Sciences) Forced Oscillations of Nonlinear Circuits. Pipes........352–8
(Basic Sciences) Harmonic Analysis for Nonlinear Characteristics. Lewis........693–9; disc. 699
(Basic Sciences) Impulse Ionization and Breakdown in Liquid Dielectrics. Sommerman, Bute, Larson........147–53; disc. 153
(Basic Sciences) Significant Measurements for Determining the Stability of High-Temperature Magnet-Wire Insulation. Scheideler.....177–81; disc. 181
(Basic Sciences) The Measurement of Electrostatic Potential Attributable to Net Ion Space Charge in Air. Carroll, Hammond........712–15; disc. 715
(Basic Sciences) The Transmission Matrix of N Alike Cascaded Networks. Storch........616–8
(Basic Sciences) The Use of Instantaneous Point Sources or Green's Functions in Evaluating Electromagnetic Fields. Smith........82–8
Basic Theory and Experimental Verification of the A-C Galvanometer. Higgins, Kneen.235–41; disc. 241
Battery Charger, A Magnetically Regulated Portable. Anderson........607–10
Bell System, Negative-Impedance Telephone Repeaters—Application in the. Rose......430–4; disc. 434
Biased Rectifier Amplifier—A Pulse Magnetic Amplifier, The. Morgan........584–90
Bibliography on Data Storage and Recording. Hollander........49–58
Blocks for Analogue Computers, Some Transistor Building. Hellerman........410–13
Bonneville Power System, A New 12-Element Automatic Oscillograph and Application on the. Hathaway, Davis, Curtin........307–12
Breakdown in Liquid Dielectrics, Impulse Ionization and. Sommerman, Bute, Larson........147–53; disc. 153
Breakdown of Solid Insulation, Effect of Electric Discharges on the. Dakin, Philofsky, Divens........155–61; disc. 161
Bridge Magnetic Amplifier, Design Considerations of the Half-Wave. Lufcy, Woodson........220–5; disc. 225
Brillouin-Wentzel-Kramers Method, Analysis of Time-Varying Circuits by the. Pipes........93–6
Broadcast Stations, A Radio Relay Remote-Control System for Frequency-Modulation. Humphrey........128–31
Broadcasting, High-Gain Side-Firing Helical Antennas for Ultrahigh-Frequency Television. Smith........135–8
Building Blocks for Analogue Computers, Some Transistor. Hellerman........410–13
Busses Comprised of Double-Channel Conductors, Equations for Determining Current Distribution Among the Conductors of. Siegel, Higgins........489–92

C

Cable Plant Design, Simplified Transmission Engineering in Exchange. Bogan, Young........498–502; disc. 502
Cable Design for Military Carrier Telephone Systems, A New. Kingsley........506–09; disc. 509
Cable, Polyethylene-Insulated Telephone. Windeler........106–11
Cable, Type-N1 Carrier on Radio and Coaxial. Ames, Wedel........312–18
Calculation of Life Characteristics of Insulation. Whitman, Doigan........193–7; disc. 197
Calls in Nationwide Toll Dialing, Some Recent Advances in the Economy of Routing. Molnar.......670–9
Carrier on Radio and Coaxial Cable, Type-N1. Ames, Wedel........312–18
Carrier System, Application and Transmission Features of a New 12-Channel Open-Wire. Appert, Caruthers, Chaskin........18–27
Carrier System, Mechanical Aspects and Component Features of a New 12-Channel Open-Wire. Ewing, Frazee, Welling........75–81
Carrier Systems, Varistor Modulators for. Caruthers........291–8
Carrier Telephone Systems, A New Cable Design for Military. Kingsley........506–09; disc. 509
Carrier Telephone Systems, Considerations for Development of New Military. Boykin, Johnston, Bedrosian........503–06
Carrier Telephone Systems Equipment Features, New Military. Hoffmann........509–15
Carrier Telephone Systems, New Military. Huber, Miller, Schramm........515–25
Carrier Telephone Systems, Servicing Center for Short-Haul. Bonner........388–96
Cascaded Networks, The Transmission Matrix of N Alike. Storch........616–8
Cascaded Self-Saturating Magnetic Amplifier Performance, Techniques for Measuring. Kaplan, Wolff........581–4
Cell Comparator, A Simplified Standard. Miller........413–15; disc. 415

Center for Military Use, A Fully Automatic Teletypewriter Switching. Johnston, Stiles.........27–37
Center for Short-Haul Carrier Telephone Systems, Servicing. Bonner.........388–96
Characteristics of Insulation, Calculation of Life. Whitman, Doigan.........193–7; disc. 197
Characteristics of Selenium Rectifiers, Instantaneous Electrical. Pittman.........45–9
Characteristics of Several Magnetic Materials, Flux Resetting. Huhta.........111–14
Characteristics of Temperature-Sensitive Elements, Shaping of the. Keonjian, Schaffner.....396–400
Characteristics of the High-Current Argon Arc with Various Electrode Materials. Dzimianski, Jones.........665–9; disc. 669
Characteristics Pertinent to the Operation of Cores in Self-Saturating Magnetic Amplifiers, Magnetic. Roberts.........682–9; disc. 689
Charge in Air, The Measurement of Electrostatic Potential Attributable to Net Ion Space. Carroll, Hammond.........712–15; disc. 715
Charger, A Magnetically Regulated Portable Battery. Anderson.........607–10
Circuit Design Aspects, Telegraph Terminal AN/FGC-29. Boughtwood.........531–6
Circuit Models of the Nuclear Reactor, Electric. Kron.........259–65
Circuit, Saturating Transformer Reference. Evans.........610–16; disc. 616
Circuit with Nonlinear Inductance and Nonlinear Capacitance, Acceleration Plane Method for Analysis of a. Ku.........619–26
Circuit with Quiescent Current Compensation, Single-Ended Saturable Reactor. Radus.......597–603
Circuits, A Symbolic Method for Synthesis of 2-Terminal Switcing. Zeheb, Caywood.........690–3
Circuits, Analysis of Nonlinear Coupled. Ku.........626–30; disc. 630
Circuits by the Brillouin-Wentzel-Kramers Method, Analysis of Linear Time-Varying. Pipes.....93–6
Circuits, Forced Oscillations of Nonlinear. Pipes.........352–8
Circuits with Full-Wave Output and Half-Wave Control Signals, Magnetic-Amplifier. Lord.......265–70
Closed-Loop Automatic Phase Control. Ordung, Gibson, Shinn.........375–81
Coaxial Cables, Type-N1 Carrier on Radio and. Ames, Wedel.........312–18
Color Television Systems, Differential Phase and Gain Measurements in. Kelly.........565–9
Committee Report, Joint AIEE-IRE) Proposed Methods of Testing Transistors. (January 1955 C & E only).........725–40
(Communication) A Fully Automatic Teletypewriter Switching Center for Military Use. Johnston, Stiles.........27–37
(Communication) A Private Microwave Radio System for Power Company Use. Hazen, Danser, Zillis.........492–8
(Communication) Application of Toll Dialing in Pennsylvania. Charny.........400–05
(Communication) Automatic Percussion Welding. Quinlan.........561–5
Communication for Safety, Technical Considerations Relating to Marine Radiotelephone. Laube.........347–52
(Communication) Interoffice Trunking and Signaling. Nolke.........590–7
(Communication) Negative-Impedance Telephone Repeaters—Application in the Bell System. Rose.........430–4; disc. 434
(Communication) Radio Transmission of Narrow-Band Mobile Radio Systems at 40-Megacycles. Kraus, Chaney, Steelman.........302–07
(Communication) Rural Automatization in the Swedish Telephone System. Bjurel.........552–61
(Communication) Some Recent Advances in the Economy of Routing Calls in Nationwide Toll Dialing. Molnar.........670–9
(Communication) The Recognition and Identification of Symmetric Switching Functions. Caldwell.........142–6; disc. 146
(Communication) Type-N1 Carrier on Radio and Coaxial Cable. Ames, Wedel.........312–18
(Communications) Application and Transmission Features of a New 12-Channel Open-Wire Carrier System. Appert, Caruthers, Chaskin.......18–27
(Communications) Mechanical Aspects and Component Features of a New 12-Channel Open-Wire Carrier System. Ewing, Frazee, Welling.........75–81
(Communications) Some Engineering Considerations in the Design of Telephone Systems to Serve Predominantly Rural Areas. McDonough, Smith.........208–13; disc. 213
Communications Systems, Industry Co-ordination of Microwave. Nexon.........250–3
(Communications) Torque Requirements of a Radar Antenna. Mark.........15–18
Community Television Systems, Line Amplifiers for. Simons, Kirk Arbeiter.........214–19; disc. 219
Comparative Propagation Studies on 250 and 450 Megacycles. Hopner, Cushing.........318–20
Comparator, A Simplified Standard Cell. Miller.........413–15; disc. 415
Comparison of Metals and Ferrites for High-Speed Pulse Operation, A. Brown, Buck, Meuyuk.........631–4; disc. 634
Compensation, Single-Ended Saturable Reactor Circuit with Quiescent Current. Radus.........597–603
Component Features of a New 12-Channel Open-Wire Carrier System, Mechanical Aspects and. Ewing, Frazee, Welling.........75–81
Computer for Automatic Determination of System Swing Curves, An Analogue. Shen, Lisser.........475–83; disc. 483
Computer Power Supplies, Precision High-Current. Rosenstein.........405–09
Computer, Rectifier Arc-Back Study on the Analogue. Dillard, Baldwin.........198–207; disc. 207
Computer Technique Using Magnetic Amplifiers, An Analogue. Davis, Swift.........635–9; disc. 639
Computers as an Aid in Electric-Machine Design, Digital. Saunders.........189–91; disc. 191
Computers, Some Transistor Building Blocks for Analogue. Hellerman.........410–13
(Computing Devices) A Digital Data-Recorder for Dense Storage of Continuous Voltages. Hollander.........253–9
(Computing Devices) Bibliography on Data Storage and Recording. Hollander.........49–58
(Computing Devices) Magnetic Drum Recording of Digital Data. Hoagland.........381–5
(Computing Devices) Networks for Digital- to Analogue Shaft-Position Transducers. O'Neil.......456–66
Conductivity at Ultrahigh Frequencies, Measurements of Materials with High Dielectric Constant and. Schwan, Li.........603–07
Conductors, Impedance and Induced Voltage Measurements on Iron. Mackenzie.....577–80; disc. 580
Conductors of Busses Comprised of Double Channel Conductors, Equations for Determining Current Distribution Among the. Siegel, Higgins..489–92
Considerations for a New Military Radio Relay System. Ribe, Brown.........547–52
Considerations for Development of New Military Carrier Telephone Systems. Boykin, Johnston, Bedrosian.........503–06
Considerations of the Half-Wave Bridge Magnetic Amplifier, Design. Lufcy, Woodson.........220–5; disc. 225
Constant and Conductivity at Ultrahigh Frequencies, Measurements of Materials with High Dielectric. Schwan, Li.........603–07
Control, Closed-Loop Automatic Phase. Ordung, Gibson, Shinn.........375–81
Control Signals, Magnetic-Amplifier Circuits with Full-Wave Output and Half-Wave. Lord......265–70
Control System for Electrical Precipitators, An Automatic Volage. Hall.........124–7
Control Systems, Analysis and Synthesis of Sampled-Data. Jury.........332–46
Controlled Magnetic Amplifiers, Derivative-. Schnitzler.........569–76
Co-ordination of Microwave Communications Systems, Industry. Nexon.........250–3
Core Functions, Analysis of a Single-Core Magnetic Amplifier with Real Rectifier and. Frank, Rabotnick, Walker.........182–8; disc. 188
Cores in Self-Saturating Magnetic Amplifiers, Magnetic Characteristics Pertinent to the Operation of. Roberts.........682–9; disc. 689
Cores, Measurement of the Quality Factor of Inductor. Stewart.........298–302
Counters, Accurate Tachometry Methods with Electronic. Shulman.........452–5; disc. 456
Coupled Circuits, Analysis of Nonlinear. Ku.........626–30; disc. 630
(Crystals) Electrical Properties of Microcrystalline Selenium. Halverson.........38–45
Cube, Switching Functions on an n-Dimensional. Lee.........289–91
Current Compensation, Single-Ended Saturable Reactor Circuit with Quiescent. Radus.........597–603
Current Distribution Among the Conductors of Busses Comprised of Double-Channel Conductors, Equations for Determining. Siegel, Higgins....489–92
Curves, An Analogue Computer for Automatic Determination of System Swing. Shen, Lisser.........475–83; disc. 483
Cyclic Integrator—A Device for Measuring the Frequency Response of Magnetic Amplifiers, The. Dunnegan, Harnden.........358–65

D

D-C Load, Magnetic Amplifiers with Inductive. Storm.........466–74; disc. 474
D-C Magnetic Amplifier, Pulse Relaxation Amplifier—A Low-Level. Morgan, McFerran.........245–9
Data Control Systems, Analysis and Synthesis of Sampled-. Jury.........332–46
Data, Magnetic Drum Recording of Digital. Hoagland.........381–5
Data-Recorder for Dense Storage of Continuous Voltages, A Digital. Hollander.........253–9
Data Storage and Recording, Bibliography on. Hollander.........49–58
Derivative-Controlled Magnetic Amplifiers. Schnitzler.........569–76
Design Aspects, Telegraph Terminal AN/FGC-29 Circuit. Boughtwood.........531–6
Design Considerations of the Half-Wave Bridge Magnetic Amplifier. Lufcy, Woodson.220–5; disc. 225
Design, Digital Computers as an Aid in Electric-Machine. Saunders.........189–91; disc. 191
Determination of System Swing Curves, An Analogue Computer for Automatic. Shen, Lisser.........475–83; disc. 483
Determining Current Distribution Among the Conductors of Busses Comprised of Double-Channel Conductors, Equations for. Siegel, Higgins....489–92
Development of New Military Carrier Telephone Systems, Considerations for. Boykin, Johnston, Bedrosian.........503–06
Device for Measuring the Frequency Response of Magnetic Amplifiers, The Cyclic Integrator—A. Dunnegan, Harnden.........358–65
Devices in the Feedback Loop of Regulated Metallic Rectifiers, Some Applications of Semiconductor. Hamilton.........640–5
Dialing in Pennsylvania, Application of Toll. Charny.........400–05
Dialing, Some Recent Advances in the Economy of Routing Calls in Nationwide Toll. Molnar.........670–9
Dielectric Amplifier, Frequency Response of a Resonant. Penney, Sack, Wingrove.........119–24
Dielectric Constant and Conductivity at Ultrahigh Frequencies, Measurements of Materials with High. Schwan, Li.........603–07
Dielectric Materials, Thickness Gauge for. Woods.........320–3
Dielectrics, Impulse Ionization and Breakdown in Liquid. Sommerman, Bute, Larson.........147–53; disc. 153
Difference Equations, Analysis of Magnetic Amplifiers by the Use of. Johannessen........700–11; disc. 711
Differential Phase and Gain Measurements in Color Television Systems. Kelly.........565–9
Digital Computers as an Aid in Electric-Machine Design. Saunders.........189–91; disc. 191
Digital Data, Magnetic Drum Redording of. Hoagland.........381–5
Digital Data-Recorder for Dense Storage of Continuous Voltages, A. Hollander.........253–9
Digital-to-Analogue Shaft-Position Transducers, Networks for. O'Neil.........456–66
Dimensional Cube, Switching Functions on an n-. Lee.........289–91
Diode as a Reference Standard in Regulated Metallic Rectifier Circuits, The Suitability of the Silicon. Alloy Junction. Smith.........645–51; disc. 651
Direct Current, A New Standard for. Fitchen, Schwieg, Tucker.........655–8; disc. 658
Discharges on the Breakdown of Solid Insulation, Effect of Electric. Dakin, Philofsky, Divens.........155–61; disc. 161
Drum Recording of Digital Data, Magnetic. Hoagland.........381–5

E

E-Type Repeaters, Theory of. Merrill.........443–7
E-Type Telephone Repeaters—Description, Equipment, and Testing. Smethurst.........435–43
Eddy-Current Losses in a Semi-Infinite Solid due to a Nearby Alternating Current. Poritsky, Jerrard.........97–106
Eddy-Current Phenomena in Ferromagnetic Materials. McConnell.........226–33; disc. 233
Effect of Electric Discharges on the Breakdown of Solid Insulation. Dakin, Philofsky, Divens.........155–61; disc. 161
Electric Circuit Models of the Nuclear Reactor. Kron.........259–65
Electric Discharges on the Breakdown of Solid Insulation, Effect of. Dakin, Philofsky, Divens.........155–61; disc. 161
Electric-Machine Design, Digital Computers as an Aid in. Saunders.........189–91; disc. 191
Electric Networks as a Function of Frequency, Precision Measurement of Phasors in. Hermont.....365–75
Electrical Characteristics of Selenium Rectifiers, Instantaneous. Pittman.........45–9
Electrical Precipitators, An Automatic Voltage Control System for. Hall.........124–7
Electrical Properties of Microcrystalline Selenium. Halverson.........38–45
Electrical Protection of Telephone Systems. Bullard, Hays, Saunders.........385–8
Electrode Materials, Characteristics of the High Current Argon Arc with Various. Dzimianski, Jones.........665–9; disc. 669

Electromagnetic Fields, The Use of Instantaneous Point Sources or Green's Functions in Evaluating. Smith 82–8
Electronic Counters, Accurate Tachometry Methods with. Shulman........452–5; disc. 456
(Electronics) An Annular Wave Guide Rotary Joint with Wave Guide Feed. Breetz........62–6
(Electronics) An Automatic Voltage Control System for Electrical Precipitators. Hall........124–7
(Electronics) Closed-Loop Automatic Phase Control. Ordung, Gibson, Shinn........375–81
(Electronics) Shaping of the Characteristics of Temperature Sensitive Elements. Keonjian, Schaffner........396–400
Electrostatic Potential Attributable to Net Ion Space Charge in Air, The Measurement of. Carroll, Hammond........712–15; disc. 715
Elements, Shaping of the Characteristics of Temperature-Sensitive. Keonjian, Schaffner........396–400
Energy in Synchronous Machines, Flow of. Hawthorne 1–10
Engineering Considerations in the Design of Telephone Systems to Serve Predominantly Rural Areas, Some. McDonough, Smith........208–13; disc. 213
Equations, Analysis of Magnetic Amplifiers by the Use of Difference. Johannessen........700-11; disc. 711
Equations for Determining Current Distribution Among the Conductors of Busses Comprised of Double-Channel Conductors. Siegel, Higgins........489–92
Equipment and Mechanical Features of the AN/TRC-24 Radio Set. Cruser........544-7
Equipment Features, Telegraph Terminal AN/FGC-29. Cusack........536–40
Equipment for the Armed Services, High-Speed Teletypewriter. Schultheiss........88–93
Exchange Cable Plant Design, Simplified Transmission Engineering in. Bogan, Young........498–502; disc. 502
Experimental Verification of the A-C Galvanometer, Basic Theory and. Higgins, Kneen........235–41; disc. 241

F

Fast Response Magnetic Amplifiers. Hughes, Miller........69–75
Feedback, Alteration of the Dynamic Response of Magnetic Amplifiers by. Decker........658–63; disc. 663
Feedback Loop of Regulated Metallic Rectifiers, Some Applications of Semiconductor Devices in the. Hamilton........640–5
Ferrites for High-Speed Pulse Operation, A Comparison of Metals and. Brown, Buck, Menyuk........631–4; disc. 634
Ferromagnetic Materials, Eddy-Current Phenomena in. McConnell........226–33; disc. 233
Firing Helical Antennas for Ultrahigh-Frequency Television Broadcasting, High-Gain Side. Smith........135–8
Flow of Energy in Synchronous Machines. Hawthorne 1–10
Flux Resetting Characteristics of Several Magnetic Materials. Huhta........111–14
Forced Oscillations of Nonlinear Circuits. Pipes.352–8
Frequencies, The Use of Sphere Gaps at Radio. Oler........329–32
Frequency-Modulation Broadcast Stations, A Radio Relay Remote-Control System for. Humphrey........128–31
Frequency Multipliers, Magnetic. Johnson, Rauch........448–51; disc. 451
Frequency, Precision Measurement of Phasors in Electric Networks as a Function of. Hermont........365–75
Frequency Response of a Resonant Dielectric Amplifier. Penney, Sack, Wingrove........119–24
Frequency Response of Magnetic Amplifiers, The Cyclic Integrator—A Device for Measuring the. Dunnegan, Harnden........358–65
Full-Wave Output and Half-Wave Control Signals, Magnetic-Amplifier Circuits with. Lord...265–70
Functions, Analysis of a Single-Core Magnetic Amplifier with Real Rectifier and Core. Frank, Rabotnick, Walker........182–8; disc. 188
Functions on an n-Dimensional Cube, Switching. Lee........289-91
Functions, The Recognition and Identification of Symmetric Switching. Caldwell........142–6; disc. 146

G

Gain Measurements in Color Television Systems, Differential Phase and. Kelly........565–9
Galvanometer, Basic Theory and Experimental Verification of the A-C. Higgins, Kneen........235–41; disc. 241
Gaps at Radio Frequencies, The Use of Sphere. Oler........329–32
Gauge for Dielectric Materials, Thickness. Woods........320–3
Green's Functions in Evaluating Electromagnetic Fields, The Use of Instantaneous Point Sources or. Smith 82–8
Guide Rotary Joint with Wave Guide Feed, An Annular Wave. Breetz........62–6

H

Half-Wave Bridge Magnetic Amplifier, Design Considerations of the. Lufcy, Woodson...220–5; disc. 225
Half-Wave Control Signals, Magnetic-Amplifier Circuits with Full-Wave Output and. Lord........265–70
Harmonic Analysis for Nonlinear Characteristics. Lewis........693–9; disc. 699
Helical Antennas for Ultrahigh-Frequency Television Broadcasting, High-Gain Side-Firing. Smith........135–8
High-Current Computer Power Supplies, Precision. Rosenstein........405–09
High Frequencies, Thermocouple-Type Ammeters for Use at Very. McAninch........241–5; disc. 245
High-Frequency Radio Systems, A New Multichannel Teletype for Use on Long-Range. Mack Levine........540–4
High-Gain Side-Firing Helical Antennas For Ultrahigh-Frequency Television Broadcasting. Smith........135–8
High-Speed Teletypewriter Equipment for the Armed Services. Schultheiss........88–93
High-Temperature Magnet-Wire Insulation, Significant Measurements for Determining the Stability of. Scheideler........177-81; disc. 181

I

Identification of Symmetric Switching Functions, The Recognition and. Caldwell........142–6; disc. 146
Impedance and Induced Voltage Measurements on Iron Conductors. Mackenzie........577–80; disc. 580
Impedance Telephone Repeaters—Application in the Bell System, Negative-. Rose....430–4; disc. 434
Impulse Ionization and Breakdown in Liquid Dielectrics. Sommerman, Bute, Larson........147–53; disc. 153
Induced Voltage Measurements on Iron Conductors, Impedance and. Mackenzie....577–80; disc. 580
Inductance and Nonlinear Capacitance, Acceleration Plane Method for Analysis of a Circuit with Nonlinear. Ku........619–26
Inductive D-C Load, Magnetic Amplifiers with. Storm ....466–74; disc. 474
Inductor Cores, Measurement of the Quality Factor of. Stewart........298–302
Industry Co-ordination of Microwave Communications Systems. Nexon........250–3
Instantaneous Electrical Characteristics of Selenium Rectifiers. Pittman........45-9
Instantaneous Point Sources on Green's Functions in Evaluating Electromagnetic Fields, The Use of. Smith........82–8
Instrument Applications, Stabilized Power Supplies for. Amey, Krantz, Clark, Williams..484–8; disc. 488
(Instruments) A New Standard for Direct Current. Fitchen, Schwieg, Tucker........655–8; disc. 658
(Instruments) A New 12-Element Automatic Oscillograph and Application on the Bonneville Power System. Hathaway, Davis, Curtin........307–12
(Instruments) A Simplified Standard Cell Comparator. Miller........413–15; disc. 415
(Instruments) A Transducer Using a Short-Circuit Rotor. Orlando........131–4
(Instruments) A Tungsten Resistance Thermometer. Sias, Macintyre, Hansen........66–9
(Instruments) Basic Theory and Experimental Verification of the A-C Galvanometer. Higgins, Kneen........235–41; disc. 241
(Instruments) The Quasi-Peak Voltmeter. Frick........417–23; disc. 423
(Instruments) The Use of Sphere Gaps at Radio Frequencies. Oler........329–32
(Instruments) Thermocouple-Type Ammeters for Use at Very High Frequencies. McAninch........241–5; disc. 245
(Instruments) Thickness Gauge for Dielectric Materials. Woods........320–3
Insulated Telephone Cable, Polyethylene-. Windeler........106–11
Insulation, Calculation of Life Characteristics of. Whitman, Doigan........193–7; disc. 197
Insulation, Effect of Electric Discharges on the Breakdown of Solid. Dakin, Philofsky, Divens........155–61; disc. 161
Insulation, Significant Measurements for Determining the Stability of High-Temperature Magnet-Wire. Scheideler........177–81; disc. 181
Integrator—A Device for Measuring the Frequency Response of Magnetic Amplifiers, The Cyclic. Dunnegan, Harnden........358–65
Interoffice Trunking and Signaling. Nolke........590–7
Ion Space Charge in Air, The Measurement of Electrostatic Potential Attributable to Net. Carroll, Hammond........712–15; disc. 715
Ionization and Breakdown in Liquid Dielectrics, Impulse. Sommerman, Bute, Larson........147–53; disc. 153
(IRE-AIEE) Proposed Methods of Testing Transistors. (Committee Report) (January 1955 C & E only)........725–40
Iron Conductors, Impedance and Induced Voltage Measurements on. Mackenzie..577–80; disc. 580

J

Joint with Wave Guide Feed, An Annular Wave Guide Rotary. Breetz........62–6
Junction Diode as a Reference Standard in Regulated Metallic Rectifier Circuits, The Suitability of the Silicon Alloy. Smith........645–51; disc. 651

K

Klystron Power Amplifiers, An Ultrahigh-Frequency Transmitter Employing. Sayer........114–18

L

Life Characteristics of Insulation, Calculation of. Whitman, Doigan........193–7; disc. 197
Line Amplifiers for Community Television Systems. Simons, Kirk, Arbeiter........214–19; disc. 219
Linear Time-Varying Circuits by the Brillouin-Wentzel-Kramers Method, Analysis of. Pipes........93–6
Liquid Dielectrics, Impulse Ionization and Breakdown in. Sommerman, Bute, Larson..147–53; disc. 153
Load, Magnetic Amplifiers with Inductive D-C. Storm ....466–74; disc. 474
Load, The Operation of Magnetic Amplifiers with Various Types of. Finzi, Jackson:
  Part I. Load Currents for Given Angle of Firing........270–8
  Part II. Controlling the Angle of Firing........279–88
Long-Range High-Frequency Radio Systems, A New Multichannel Teletype for Use on. Mack, Levine........540–4
Losses in a Semi-Infinite Solid due to a Nearby Alternating Current, Eddy-Current. Poritsky, Jerrard........97–106
Low-Level D-C Magnetic Amplifier, Pulse Relaxation Amplifier—A. Morgan, McFerran........245–9

M

Machine Design, Digital Computers as an Aid in Electric. Saunders........189–91; disc. 191
Machines, Flow of Energy in Synchronous. Hawthorne 1–10
Magnet-Wire Insulation, Significant Measurements for Determining the Stability of High-Temperature. Scheideler........177–81; disc. 181
Magnetic-Amplifier Circuits with Full-Wave Output and Half-Wave Control Signals. Lord........265–70
Magnetic Amplifier, Design Considerations of the Half-Wave Bridge. Lufcy, Woodson....220–5; disc. 225
Magnetic Amplifier for Airborne Servo-Control Systems, An Operational-Type. Hubbard.425–30; disc. 430
Magnetic-Amplifier Performance, Techniques for Measuring Cascaded Self-Saturating. Kaplan, Wolff........581–4
Magnetic Amplifier, Pulse Relaxation Amplifier—A Low-Level D-C. Morgan, McFerran........245–9
Magnetic Amplifier, The Biased Rectifier Amplifier—A Pulse. Morgan........584–90
Magnetic Amplifier with Real Rectifier and Core Functions, Analysis of a Single-Core. Frank, Rabotnick, Walker........182–8; disc. 188
Magnetic Amplifiers, An Analogue Computer Technique Using. Davis, Swift........635–9; disc. 639
Magnetic Amplifiers by Feedback, Alteration of the Dynamic Response of. Decker........658–63; disc. 663
Magnetic Amplifiers by the Use of Difference Equations, Analysis of. Johannessen........700–11; disc. 711
Magnetic Amplifiers, Derivative-Controlled. Schnitzler 569–76
Magnetic Amplifiers, Fast-Response. Hughes, Miller........69–75
(Magnetic Amplifiers) Frequency Response of a Resonant Dielectric Amplifier. Penney, Sack, Wingrove 119–24
Magnetic Amplifiers, Magnetic Characteristics Pertinent to the Operation of Cores in Self-Saturating. Roberts........682–9; disc. 689
(Magnetic Amplifiers) Saturating Transformer Reference Circuit. Evans........610–16; disc. 616
(Magnetic Amplifiers) Single-Ended Saturable Reactor Circuit with Quiescent Current Compensation. Radus........597–603

Magnetic Amplifiers with Various Types of Load, The Operation of. Finzi, Jackson:
  Part I. Load Currents for Given Angle of Firing........270–8
  Part II. Controlling the Angle of Firing.....279–88
Magnetic Amplifiers, The Cyclic Integrator—A Device for Measuring the Frequency Response of. Dunnegan, Harnden........358–65
Magnetic Amplifiers with Inductive D-C Load. Storm........466–74; disc. 474
Magnetic Characteristics Pertinent to the Operation of Cores in Self-Saturating Magnetic Amplifiers, Roberts........682–9; disc. 689
Magnetic Drum Recording of Digital Data. Hoagland........381–5
Magnetic Frequency Multipliers. Johnson, Rauch........448–51; disc. 451
Magnetic Materials, Flux Resetting Characteristics of Several. Huhta........111–14
Magnetic Recording, A Survey of. Begun........651–5; disc. 655
Magnetically Regulated Portable Battery Charger. A Anderson........607–10
Marine Radiotelephone Communication for Safety, Technical Considerations Relating to. Laube........347–52
Materials, Eddy-Current Phenomena in Ferromagnetic. McConnell........226–33; disc. 233
Materials, Flux Resetting Characteristics of Several Magnetic. Huhta........111–14
Materials, Thickness Gauge for Dielectric. Woods........320–3
Materials with High Dielectric Constant and Conductivity at Ultrahigh Frequencies, Measurements of. Schwan, Li........603–07
Matrix of N Alike Cascaded Networks, The Transmission. Storch........616–8
Measurement of Electrostatic Potential Attributable to Net Ion Space Charge in Air, The. Carroll, Hammond........712–15; disc. 715
Measurement of Phasors in Electric Networks as a Function of Frequency, Precision. Hermont....365–75
(Measurements) Accurate Tachometry Methods with Electronic Counters. Shulman...452–5; disc. 456
Measurements for Determining the Stability of High-Temperature Magnet-Wire Insulation, Significant. Scheideler........177–81; disc. 181
Measurements in Color Television Systems, Differential Phase and Gain. Kelly........565–9
Measurements of Materials with High Dielectric Constant and Conductivity at Ultrahigh Frequencies. Schwan, Li........603–07
Measurements of the Quality Factor of Inductor Cores. Stewart........298–302
Measurements on Iron Conductors, Impedance and Induced Voltage. Mackenzie.....577–80; disc. 580
Measurements with a Phase-Angle-Type Probe, Sheet and Plated-Metal. Yates, Queen........138–42
Measuring Set, A New Portable Telegraph Transmission. Cory........59–62
Measuring the Frequency Response of Magnetic Amplifiers, The Cyclic Integrator—A Device for. Dunnegan, Harnden........358–65
Mechanical Aspects and Component Features of a New 12-Channel Open-Wire Carrier System. Ewing, Frazee, Welling........75–81
Mechanical Features of the AN/TRC-24 Radio Set, Equipment and. Cruser........544–7
Megacycles, Comparative Propagation Studies on 250 and 250. Hopner, Cushing........318–20
Megacycles, Radio Transmission of Narrow-Band Mobile Radio Systems at 40. Kraus, Chaney, Steelman........403–07
Metal Measurements with a Phase-Angle-Type Probe, Sheet and Plated-. Yates, Queen........138–42
Metallic Rectifier Circuits, The Suitability of the Silicon Alloy Junction Diode as a Reference Standard in Regulated. Smith........645–51; disc. 651
(Metallic Rectifiers) A Magnetically Regulated Portable Battery Charger. Anderson........607–10
(Metallic Rectifiers) Electrical Properties of Microcrystalline Selenium. Halverson........38–45
Metallic Rectifiers, Some Application of Semiconductor Devices in the Feedback Loop of Regulated. Hamilton........640–5
Metals and Ferrites for High-Speed Pulse Operation, A Comparison of. Brown, Buck, Menyuk........631–4; disc. 634
Method for Analysis of a Circuit with Nonlinear Inductance and Nonlinear Capacitance, Acceleration Plane. Ku........619–26
Method for Synthesis of 2-Terminal Switching Circuits, A Symbolic. Zeheb, Caywood........690–3
Methods of Testing Transistors, Proposed. (Joint AIEE-IRE Committee Report). (January 1955 C & E only)........725–40
Methods with Electronic Counters, Accurate Tachometry. Shulman........452–5; disc. 456
Microcrystalline Selenium, Electrical Properties of. Havlerson........38–45
Microwave Communications Systems, Industry Coordination of. Nexon........250–3
Microwave Radio System for Power Company Use, A Private. Hazen, Danser, Zillis........492–8
Military Carrier Telephone Systems, A New Cable Design for. Kingsley........506–09; disc. 509
Military Carrier Telephone Systems, Considerations for Development of New. Boykin, Johnston, Bedrosian........503–06
Military Carrrier Telephone Systems Equipment Features, New. Hoffmann........509–15
Military Carrier Telephone Systems, New. Huber, Miller, Schramm........515–25
(Military) High-Speed Teletypewriter Equipment for the Armed Services. Schultheiss........88–93
Military Radio Relay System, A New Ultrahigh-Frequency Multichannel. Nordahl........526–31
Military Radio Relay System, Considerations for a New. Ribe, Brown........547–52
Military Use, A Fully Automatic Teletypewriter Switching Center for. Johnston, Stiles........27–37
Mobile Radio Systems at 40 Megacycles, Radio Transmission of Narrow-Band. Kraus, Chaney, Steelman........403–07
Models of the Nuclear Reactor, Electric Circuit. Kron........259–65
Modulators for Carrier Systems, Varistor. Caruthers........291–8
Multichannel Military Radio Relay System, A New Ultrahigh-Frequency. Nordahl........526–31
Multichannel Teletype for Use on Long-Range High-Frequency Radio Systems, A New. Mack, Levine........540–4
Multipliers, Magnetic Frequency. Johnson, Rauch........448–51; disc. 451

N

n-Dimensional Cube, Switching Functions on an. Lee........289–91
Narrow-Band Mobile Radio Systems at 40-Megacycles, Radio Transmission of. Kraus, Chaney, Steelman........403–07
Nationwide Toll Dialing, Some Recent Advances in the Economy of Routing Calls in. Molnar......670–9
Negative-Impedance Telephone Repeaters—Application in the Bell System. Rose........430–4; disc. 434
Network Theory to the Analysis of Rotating Machinery, Application of. Koenig:
  Part I. Synchronous and Asynchronous Machines........162–9
  Part II. Commutating Machines........169–74
  Discussions........174
Networks as a Function of Frequency, Precision Measurement of Phasors in Electric. Hermont.....365–75
Networks for Digital-to-Analogue Shaft-Position Transducers. O'Neil........456–66
Networks, The Transmission Matrix of N Alike Cascaded. Storch........616–8
Nonlinear Characteristics, Harmonic Analysis for. Lewis........693–9; disc. 699
Nonlinear Circuits, Forced Oscillations of. Pipes..352–8
Nonlinear Coupled Circuits, Analysis of. Ku........626–30; disc. 630
Nonlinear Inductance and Nonlinear Capacitance, Acceleration Plane Method for Analysis of a Circuit with. Ku........619–26
Nuclear Reactor, Electric Circuit Models of the. Kron........259–65

O

One Approach to a Video SHF Relay System. Coe, McClatchie........323–9
Open-Wire Carrier System, Application and Transmission Features of a New 12-Channel. Appert, Caruthers, Chaskin........18–27
Open-Wire Carrier System, Mechanical Aspects and Component Features of a New 12-Channel. Ewing, Frazee, Welling........75–81
Operation of Magnetic Amplifiers with Various Types of Load, The. Finzi, Jackson:
  Part I. Load Currents for Given Angle of Firing........270–8
  Part II. Controlling the Angle of Firing.....279–88
Operational-Type Magnetic Amplifier for Airborne Servo-Control Systems, An. Hubbard........425–30; disc. 430
Oscillations of Nonlinear Circuits, Forced. Pipes..352–8
Oscillograph and Application on the Bonneville Power System, A New 12-Element Automatic. Hathaway, Davis, Curtin........307–12
Output and Half-Wave Control Signals, Magnetic-Amplifier Circuits with Full-Wave. Lord..265–70

P

Peak Voltmeter, The Quasi-. Frick..417–23; disc. 423
Pennsylvania, Application of Toll Dialing in. Charny........400–05
Percussion Welding, Automatic. Quinlan.......561–5
Phase and Gain Measurements in Color Television Systems, Differential. Kelly........565–9
Phase-Angle-Type Probe, Sheet and Plated-Metal Measurements with a. Yates, Queen......138–42
Phase Control, Closed-Loop Automatic. Ordung, Gibson, Shinn........375–81
Phasors in Electric Networks as a Function of Frequency, Precision Measurement of. Hermont......365–75
Phenomena in Ferromagnetic Materials, Eddy-Current. McConnell........226–33; disc. 233
Plane Method for Analysis of a Circuit with Nonlinear Inductance and Nonlinear Capacitance, Acceleration. Ku........619–26
Plant Design, Simplified Transmission Engineering in Exchange Cable. Bogan, Young..498–502; disc. 502
Plated-Metal Measurements with a Phase-Angle-Type Proge, Sheet and. Yates, Queen........138–42
Point Sources or Green's Functions in Evaluating Electromagnetic Fields, The Use of Instantaneous. Smith........82–8
Polyethylene-Insulated Telephone Cable. Windeler........106–11
Portable Battery Charger, A Magnetically Regulated. Anderson........607–10
Portable Telegraph Transmission Measuring Set, A New. Cory........59–62
Potential Attributable to Net Ion Space Charge in Air, The Measurement of Electrostatic. Carroll, Hammons........712–15; disc. 715
Power Amplifiers, An Ultrahigh-Frequency Transmitter Employing Klystron. Sayer........114–18
Power Company Use, A Private Microwave Radio System for. Hazen, Danser, Zillis........492–8
Power Supplies for Instrument Applications, Stabilized. Amey, Krantz, Clark, Williams...484–8; disc. 488
Power Supplies, Precision High-Current Computer. Rosenstein........405–09
Precipitators, An Automatic Voltage Control System for Electrical. Hall........124–7
Precision High-Current Computer Power Supplies. Rosenstein........405–09
Precision Measurement of Phasors in Electric Networks as a Function of Frequency. Hermont.....365–75
Printing Telegraphy, A Step Forward in. Benjamin, Zenner........10–15
Private Microwave Radio System for Power Company Use, A. Hazen, Danser, Zillis........492–8
Probe, Sheet and Plated-Metal Measurements with a Phase-Angle-Type. Yates, Queen........138–42
Propagation Studies on 250 and 450 Megacycles, Comparative. Hopner, Cushing........318–20
Properties of Microcrystalline Selenium, Electrical. Halverson........38–45
Protection of Telephone Systems, Electrical. Bullard, Hays, Saunders........385–8
Pulse Magnetic Amplifier, The Biased Rectifier Amplifier—A. Morgan........584–90
Pulse Operation, A Comparison of Metals and Ferrites for High-Speed. Brown, Buck, Menyuk........631–4; disc. 634
Pulse Relaxation Amplifier—A Low-Level D-C Magnetic Amplifier. Morgan, McFerran.......245–9

Q

Quality Factor of Inductor Cores, Measurement of the. Stewart........298–302
Quasi-Peak Voltmeter, The. Frick....417–23; disc. 423
Quiescent Current Compensation, Single-Ended Saturable Reactor Circuit with. Radus........597–603

R

Radar Antenna, Torque Requirements of a. Mark........15–18
Radio and Coaxial Cable, Type-N1 Carrier on. Ames, Wedel........312–18
(Radio Communications) Telegraph Terminal AN/-FGC-29 Circuit Design Aspects. Boughtwood.531–6
(Radio Communication) Telegraph Terminal AN/-FGC-29 Equipment Features. Cusack.....536–40
(Radio) Comparative Propagation Studies on 250 and 450 Megacycles. Hopner, Cushing........318–20
Radio Frequencies, The Use of Sphere Gaps at. Oler........329–32
(Radio) Industry Co-ordination of Microwave Communications Systems. Nexon........250–3
(Radio) One Approach to a Video SHF Relay System. Coe, McClatchie........323–9
Radio Relay Remote-Control System for Frequency-Modulation Broadcast Stations, A. Humphrey........128–31
Radio Relay System, A New Ultrahigh-Frequency Multichannel Military. Nordahl........526–31
Radio Relay System, Considerations for a New Military. Ribe, Brown........547–52
Radio Set, Equipment and Mechanical Features of the AN/TRC-24. Cruser........544–7

Radio System for Power Company Use, A Private Microwave. Hazen, Danser, Zillis....492–8
Radio Systems, A New Multichannel Teletype for Use on Long-Range High-Frequency. Mack, Levine....540–4
Radio Transmission of Narrow-Band Mobile Radio Systems at 40 Megacycles. Kraus, Chaney, Steelman....403–07
Radiotelephone Communication for Safety, Technical Considerations Relating to Marine. Laube.347–52
Reactor Circuit with Quiescent Current Compensation, Single-Ended Saturable. Radus....597–603
Reactor, Electric Circuit Models of the Nuclear. Kron....259–65
Recognition and Identification of Symmetric Switching Functions, The. Caldwell....142–6; disc. 146
Recorder for Dense Storage of Continuous Voltages, A Digital Data-. Hollander....253–9
Recording, A Survey of Magnetic. Begun....651–5; disc. 655
Recording, Bibliography on Data Storage and. Hollander....49–58
Recording of Digital Data, Magnetic Drum. Hoagland....381–5
Rectifier Algebra, A. Schaefer....679–82
Rectifier Amplifier—A Pulse Magnetic Amplifier, The Biased. Morgan....584–90
Rectifier and Core Functions, Analysis of a Single-Core Magnetic Amplifier with Real. Frank, Rabotnick, Walker....182–8; disc. 188
Rectifier Arc-Back Study on the Analogue Computer. Dillard, Baldwin....198–207; disc. 207
Rectifiers, Instantaneous Electrical Characteristics of Selenium. Pittman....45–9
Reference Circuit, Saturating Transformer. Evans....610–16; disc. 616
Regulated Portable Battery Charger, A Magnetically. Anderson....607–10
Relaxation Amplifier—A Low-Level D-C Magnetic Amplifier, Pulse. Morgan, McFerran....245–9
Relay Remote-Control System for Frequency-Modulation Broadcast Stations, A Radio. Humphrey....128–31
Relay System, A New Ultrahigh-Frequency Multichannel Military Radio. Nordahl....526–31
Relay System, Considerations for a New Military Radio. Ribe, Brown....547–52
Relay System, One Approach to a Video SHF. Coe, McClatchie....323–9
Remote-Control System for Frequency-Modulation Broadcast Stations, A Radio Relay. Humphrey....128–31
Repeaters—Application in the Bell System, Negative-Impedance Telephone. Rose....430–4; disc. 434
Repeaters—Description, Equipment, and Testing, E-Type Telephone. Smethurst....435–43
Repeaters, Theory of E-Type. Merrill....443–7
Requirements of a Radar Antenna, Torque. Mark....15–18
Resetting Characteristics of Several Magnetic Materials, Flux. Huhta....111–14
Resistance Thermometer, A Tungsten. Sias, Macintyre, Hansen....66–9
Resonant Dielectric Amplifier, Frequency Response of a. Penney, Sack, Wingrove....119–24
Response of a Resonant Dielectric Amplifier, Frequency. Penney, Sack, Wingrove....119–24
Response of Magnetic Amplifiers by Feedback, Alteration of the Dynamic. Decker....658–63; disc. 663
Rotating Machinery, Application of Network Theory to the Analysis of. Koenig:
  Part I. Synchronous and Asynchronous Machines....162–9
  Part II. Commutating Machines....169–74
  Discussions....174
Rotary Joint with Wave Guide Feed, An Annular Wave Guide. Breetz....62–6
Rotor, A Transducer Using a Short-Circuit. Orlando....131–4
Routing Calls in Nationwide Toll Dialing, Some Recent Advances in the Economy of. Molnar....670–9
Rural Areas, Some Engineering Considerations in the Design of Telephone Systems to Serve Predominantly. McDonough, Smith....208–13; disc. 213
Rural Automatization in the Swedish Telephone System. Bjurel....552–61

S

Safety, Technical Considerations Relating to Marine Radiotelephone Communication for. Laube....347–52
Sampled-Data Control Systems, Analysis and Synthesis of. Jury....332–46
Saturable Reactor Circuit with Quiescent Current Compensation, Single-Ended. Radus....597–603
Saturating Magnetic-Amplifier Performance, Techniques for Measuring Cascaded Self-. Kaplan, Wolff....581–4
Saturating Transformer Reference Circuit. Evans....610–16; disc. 616
Selenium, Electrical Properties of Microcrystalline. Halverson....38–45
Selenium Rectifiers, Instantaneous Electrical Characteristics of. Pittman....45–9
Self-Saturating Magnetic Amplifier Performance, Techniques for Measuring Cascaded. Kaplan, Wolff....581–4
Self-Saturating Magnetic Amplifiers, Magnetic Characteristics Pertinent to the Operation of Cores in. Roberts....682–9; disc. 689
Semiconductor Devices in the Feedback Loop of Regulated Metallic Rectifiers, Some Applications of. Hamilton....640–5
Sensitive Elements, Shaping of the Characteristics of Temperature-. Keonjian, Schaffner....396–400
Servicing Center for Short-Haul Carrier Telephone Systems. Bonner....388–96
Servo-Control Systems, An Operational-Type Magnetic Amplifier for Airborne. Hubbard....425–30; disc. 430
Shaft-Position Transducers, Networks for Digital-to-Analogue. O'Neil....456–66
Shaping of the Characteristics of Temperature-Sensitive Elements. Keonjian, Schaffner....396–400
Sheet and Plated-Metal Measurements with a Phase-Angle-Type Probe. Yates, Queen....138–42
SHF Relay System, One Approach to a Video. Coe, McClatchie....323–9
Short-Circuit Rotor, A Transducer Using a. Orlando....131–4
Short-Haul Carrier Telephone Systems, Servicing Center for. Bonner....388–96
Side-Firing Helical Antennas for Ultrahigh-Frequency Television Broadcasting, High-Gain. Smith..135–8
Signaling, Interoffice Trunking and. Nolke....590–7
Signals, Magnetic-Amplifier Circuits with Full-Wave Output and Half-Wave Control. Lord....265–70
Significant Measurements for Determining the Stability of High-Temperature Magnet-Wire Insulation. Scheideler....177–81; disc. 181
Silicon Alloy Junction Diode as a Reference Standard in Regulated Metallic Rectifier Circuits, The Suitability of the. Smith....645–51; disc. 651
Simplified Standard Cell Comparator, A. Miller....413–15; disc. 415
Simplified Transmission Engineering in Exchange Cable Plant Design. Bogan, Young....498–502; disc. 502
Single-Core Magnetic Amplifier with Real Rectifier and Core Functions, Analysis of a. Frank, Rabotnick, Walker....182–8; disc. 188
Single-Ended Saturable Reactor Circuit with Quiescent Current Compensation. Radus....597–603
Solid due to a Nearby Alternating Current, Eddy-Current Losses in a Semi-Infinite. Poritsky, Jerrard....97–106
Solid Insulation, Effect of Electric Discharges on the Breakdown of. Dakin, Philofsky, Divens....155–61; disc. 161
Some Engineering Considerations in the Design of Telephone Systems to Serve Predominantly Rural Areas. McDonough, Smith....208–13; disc. 213
Some Transistor Building Blocks for Analogue Computers. Hellerman....410–13
Space Charge in Air, The Measurement of Electrostatic Potential Attributable to Net Ion. Carroll, Hammond....712–15; disc. 715
Sphere Gaps at Radio Frequencies, The Use of. Oler....329–32
Stability of High-Temperature Magnet-Wire Insulation, Significant Measurements for Determining the. Scheideler....177–81; disc. 181
Stabilized Power Supplies for Instrument Applications. Amey, Krantz, Clark, Williams...484–8; disc. 488
Standard Cell Comparator, A Simplified. Miller....413–15; disc. 415
Standard in Regulated Metallic Rectifier Circuits, The Suitability of the Silicon Alloy Junction Diode as a Reference. Smith....645–51; disc. 651
Standard for Direct Current, A New. Fitchen, Schwieg, Tucker....655–8; disc. 658
Step Forward in Printing Telegraphy, A. Benjamin, Zenner....10–15
Storage of Continuous Voltages, A Digital Data-Recorder for Dense. Hollander....253–9
Studies on 250 and 450 Megacycles, Comparative Propagation. Hopner, Cushing....318–20
Suitability of the Silicon Alloy Junction Diode as a Reference Standard in Regulated Metallic Rectifier Circuits, The. Smith....645–51; disc. 651
Survey of Magnetic Recording, A. Begun....651–5; disc. 655
Swedish Telephone System, Rural Automatization in the. Bjurel....552–61
Swing Curves, An Analogue Computer for Automatic Determination of System. Shen, Lisser....475–83; disc. 483
Switching Center for Military Use, A Fully Automatic Teletypewriter. Johnston, Stiles....27–37
Switching Circuits, A Symbolic Method for Synthesis of 2-Terminal. Zeheb, Caywood....690–3
Switching Functions on an n-Dimensional Cube. Lee....289–91
Switching Functions, The Recognition and Identification of Symmetric. Caldwell....142–6; disc. 146
Symbolic Method for Synthesis of 2-Terminal Switching Circuits, A. Zeheb, Caywood....690–3
Symmetric Switching Functions, The Recognition and Identification of. Caldwell....142–6; disc. 146
Synchronous Machines, Flow of Energy in. Hawthorne....1–10
Synthesis of Sampled-Data Control Systems, Analysis and. Jury....332–46
Synthesis of 2-Terminal Switching Circuits, A Symbolic Method for. Zeheb, Caywood....690–3
System for Electrical Precipitators, An Automatic Voltage Control. Hall....124–7
System Swing Curves, An Analogue Computer for Automatic Determination of. Shen, Lisser....475–83; disc. 483

T

Tachometry Methods with Electronic Counters, Accurate. Shulman....452–5; disc. 456
Technical Considerations Relating to Marine Radiotelephone Communication for Safety. Laube....347–52
Techniques for Measuring Cascaded Self-Saturating Magnetic-Amplifier Performance. Kaplan, Wolff....581–4
(Telegraph) High Speed Teletypewriter Equipment for the Armed Services. Schultheiss....88–93
Telegraph Terminal AN/FGC-29 Circuit Design Aspects. Boughtwood....531–6
Telegraph Terminal AN/FGC-29 Equipment Features. Cusack....536–40
Telegraph Transmission Measuring Set, A New Portable. Cory....59–62
Telegraphy, A Step Forward in Printing. Benjamin, Zenner....10–15
Telephone Cable, Polyethylene-Insulated. Windeler....106–11
Telephone Repeaters—Application in the Bell System, Negative-Impedance. Rose....430–4; disc. 434
Telephone Repeaters—Description, Equipment, and Testing, E-Type. Smethurst....435–43
Telephone System, Rural Automatization in the Swedish. Bjurel....552–61
Telephone Systems, A New Cable Design for Military Carrier. Kingsley....506–09; disc. 509
Telephone Systems, Considerations for Development of New Military Carrier. Boykin, Johnston, Bedrosian....503–06
Telephone Systems, Electrical Protection of. Bullard, Hays, Saunders....385–8
Telephone Systems Equipment Features, New Military Carrier. Hoffmann....509–15
Telephone Systems, New Military Carrier. Huber, Miller, Schramm....515–25
Telephone Systems, Servicing Center for Short-Haul Carrier. Bonner....388–96
Telephone Systems to Serve Predominantly Rural Areas, Some Engineering Considerations in the Design of. McDonough, Smith....208–13; disc. 213
Teletype for Use on Long-Range High-Frequency Radio Systems, A New Multichannel. Mack, Levine....540–4
Teletypewriter Equipment for the Armed Services, High-Speed. Schultheiss....88–93
Teletypewriter Switching Center for Military Use, A Fully Automatic. Johnston, Stiles....27–37
(Television) A Radio Relay Remote-Control System for Frequency-Modulation Broadcast Stations. Humphrey....128–31
(Television) A Survey of Magnetic Recording. Begun....651–5; disc. 655
(Television) An Ultrahigh-Frequency Transmitter Employing Klystron Power Amplifiers. Sayer..114–18
Television Broadcasting, High-Gain Side-Firing Helical Antennas for Ultrahigh-Frequency. Smith..135–8
Television Systems, Differential Phase and Gain Measurements in Color. Kelly....565–9
Television Systems, Line Amplifiers for Community. Simons, Kirk, Arbeiter....214–19; disc. 219
Temperature-Sensitive Elements, Shaping of the Characteristics of. Keonjian, Schaffner....396–400
Terminal AN/FGC-29 Circuit Design Aspects, Telegraph. Boughtwood....531–6
Terminal AN/FGC-29 Equipment Features, Telegraph. Cusack....536–40
Testing Transistors, Proposed Methods of. (Joint AIEE-IRE Committee Report.) (January 1955 C & E only)....725–40
Theory and Experimental Verification of the A-C Galvanometer, Basic. Higgins, Kneen....235–41; disc. 241
Theory of E-Type Repeaters. Merrill....443–7
Theory to the Analysis of Rotating Machinery, Application of Network. Koenig:
  Part I. Synchronous and Asynchronous Machines....162–9
  Part II. Commutating Machines....169–74
  Discussions....174

Thermocouple-Type Ammeters for Use at Very High Frequencies. McAninch.........241–5; disc. 245
Thermometer, A Tungsten Resistance. Sias, Macintyre, Hansen..................................66–9
Thickness Gauge for Dielectric Materials. Woods..320–3
Time-Varying Circuits by the Brillouin-Wentzel-Kramers Method, Analysis of Linear. Pipes..93–6
Toll Dialing in Pennsylvania, Application of. Charny..400–05
Toll Dialing, Some Recent Advances in the Economy of Routing Calls in Nationwide. Molnar......670–9
Torque Requirements of a Radar Antenna. Mark..15–18
Transducer Using a Short-Circuit Rotor, A. Orlando..131–4
Transducers, Networks for Digital-to-Analogue Shaft-Position. O'Neil..........................456–66
Transformer Reference Circuit, Saturating. Evans..610–16; disc. 616
Transistor Building Blocks for Analogue Computers, Some. Hellerman..........................410–13
Transistors, Proposed Methods of Testing. (Joint AIEE-IRE Committee Report.) (January 1955 C & E only)..........................725–40
Transmission Engineering in Exchange Cable Plant Design, Simplified. Bogan, Young..498–502; disc. 502
Transmission Features of a New 12-Channel Open-Wire Carrier System, Application and. Appert, Caruthers, Chaskin..........................18–27
Transmission Matrix of N Alike Cascaded Networks, The. Storch..........................616–18
Transmission Measuring Set, A New Portable Telegraph. Cory..........................59–62
Transmission of Narrow-Band Mobile Radio Systems at 40 Megacycles, Radio. Kraus, Chaney, Steelman..........................403–07
Transmitter Employing Klystron Power Amplifiers, An Ultrahigh-Frequency. Sayer..............114–18
Trunking and Signaling, Interoffice. Nolke......590–7
Tungsten Resistance Thermometer, A. Sias, Macintyre, Hansen..........................66–9
Twelve-Channel Open-Wire Carrier System, Application and Transmission Features of a New. Appert, Caruthers, Chaskin..........................18–27
Twelve-Channel Open-Wire Carrier System, Mechanical Aspects and Component Features of a New. Ewing, Frazee, Welling..........................75–81
Twelve-Element Automatic Oscillograph and Application on the Bonneville Power System, A New. Hathaway, Davis, Curtin..........................307–12
250 and 450 Megacycles, Comparative Propagation Studies on. Hopner, Cushing..............318–20
Two-Terminal Switching Circuits, A Symbolic Method for Synthesis of. Zeheb, Caywood.........690–3
Type-N1 Carrier on Radio and Coaxial Cable. Ames, Wedel..........................312–18

U

Ultrahigh Frequencies, Measurements of Materials with High Dielectric Constant and Conductivity at. Schwan, Li..........................603–07
Ultrahigh-Frequency Multichannel Military Radio Relay Systems, A New. Nordahl..............526–31
Ultrahigh-Frequency Television Broadcasting, High-Gain Side-Firing Helical Antennas for. Smith..135–8
Ultrahigh-Frequency Transmitter Employing Klystron Power Amplifiers, An. Sayer..............114–18
Use of Instantaneous Point Sources or Green's Functions in Evaluating Electromagnetic Fields, The. Smith..82–8
Use of Sphere Gaps at Radio Frequencies, The. Oler..329–32

V

Varistor Modulators for Carrier Systems. Caruthers..291–8
Video SHF Relay System, One Approach to a. Coe, McClatchie..........................323–9
Voltage Control System for Electrical Precipitators, An Automatic. Hall..........................124–7
Voltage Measurements on Iron Conductors, Impedance and Induced. Mackenzie.......577–80; disc. 580
Voltages, A Digital Data-Recorder for Dense Storage of Continuous. Hollander..........................253–9
Voltmeter, The Quasi-Peak. Frick....417–23; disc. 423

W

Wave Guide Rotary Joint with Wave Guide Feed, An Annular. Breetz..........................62–6
Wave Output and Half-Wave Control Signals, Magnetic-Amplifier Circuits with Full-. Lord..265–70
Welding, Automatic Percussion. Quinlan.......561–5
(Wire Communications) A New Cable Design for Military Carrier Telephone Systems. Kingsley..506–09; disc. 509
(Wire Communications) Considerations for Development of New Military Carrier Telephone Systems. Boykin, Johnston, Bedrosian..............503–06
(Wire Communications) E-Type Telephone Repeaters—Description, Equipment, and Testing. Smethurst..435–43
(Wire Communications) Electrical Protection of Telephone Systems. Bullard, Hays, Saunders....385–8
(Wire Communications) Line Amplifiers for Community Television Systems. Simons, Kirk, Arbeiter..214–19; disc. 219
(Wire Communications) New Military Carrier Telephone Systems. Huber, Miller, Schramm..515–25
(Wire Communications) New Military Carrier Telephone Systems Equipment Features. Hoffmann..509–15
(Wire Communications) Polyethylene-Insulated Telephone Cable. Windeler..............106–11
(Wire Communications) Servicing Center for Short-Haul Carrier Telephone Systems. Bonner......388–96
(Wire Communications) Simplified Transmission Engineering in Exchange Cable Plant Design. Bogan, Young..........................498–502; disc. 502
(Wire Communications) Theory of E-Type Repeaters. Merrill..........................433–7
(Wire Communications) Varistor Modulators for Carrier Systems. Caruthers..........................291–8

2. Author Index

Ames, W. S.; W. H. Wedel. Type-N1 Carrier on Radio and Coaxial Cable. (54-214).......312–18
Amey, W. G.; F. H. Krantz, W. R. Clark, A. J. Williams, Jr. Stabilized Power Supplies for Instrument Applications. (54-295).....484–8; disc. 489
Anderson, J. G.; T. W. Liao. Disc..............153
Anderson, R. E. D. A Magnetically Regulated Portable Battery Charger. (54-416)..........607–10
Appert, K. E.; R. S. Caruthers, W. S. Chaskin. Application and Transmission Features of a New 12-Channel Open-Wire Carrier System. (54-58)..18–27
Arbeiter, H. J.; K. A. Simons, D. Kirk. Line Amplifiers for Community Television Systems. (54-101)..........................214–19; disc. 220

B

Baldwin, C. J., Jr.; J. K. Dillard. Rectifier Arc-Back Study on the Analogue Computer. (54-43)..........................198–207; disc. 207
Bedrosian, S. D.; R. S. Boykin, J. H. Johnston. Considerations for Development of New Military Carrier Telephone Systems. (54-408).........503–06
Begun, S. J. A Survey of Magnetic Recording. (54-383)..........................651–5; disc. 655
Bekey, G. A. Disc..........................483
Benjamin, A. S.; W. J. Zenner. A Step Forward in Printing Telegraphy. (54-115)............10–15
Bjurel, Bertil. Rural Automatization in the Swedish Telephone System. (54-424)..............552–61
Bogan, L. B.; K. D. Young. Simplified Transmission Engineering in Exchange Cable Plant Design. (54-411)..........................498–502
Bonner, A. L. Servicing Center for Short-Haul Carrier Telephone Systems. (54-288)............338–96
Borden, Perry A. Disc..........................241
Boughtwood, J. E. Telegraph Terminal AN/FGC-29 Circuit Design Aspects. (54-403)..........531–6
Boyajian, Clement L. Disc..........................689
Boykin, R. S.; J. H. Johnston, S. D. Bedrosian. Considerations for Development of New Military Carrier Telephone Systems. (54-408).........503–06
Breetz, Louis D. An Annular Wave Guide Rotary Joint with Wave Guide Feed. (54-179)......62–6
Brooks, H. B. Disc..........................415
Brown, D. R.; D. A. Buck, N. Menyuk. A Comparison of Metals and Ferrites for High-Speed Pulse Operation. (54-361)................631–4; disc. 634
Brown, S. P.; M. L. Ribe. Considerations for a New Military Radio Relay System. (54-404)...547–52
Buck, D. A.; D. R. Brown, N. Menyuk. A Comparison of Metals and Ferrites for High-Speed Pulse Operation. (54-361)................631–4; disc. 634
Bullard, W. R.; J. B. Hays, H. O. Saunders. Electrical Protection of Telephone Systems. (54-289)..385–8
Bute, C. J.; G. M. L. Sommerman, E. L. C. Larson. Impulse Ionization and Breakdown in Liquid Dielectrics. (54-69).............147–53; disc. 154

C

Caldwell, S. H. The Recognition and Identification of Symmetric Switching Functions. (54-24)..........................142–6; disc. 147
Carroll, J. S.; S. B. Hammond. The Measurement of Electrostatic Potential Attributable to Net Ion Space Charge in Air. (54-511)..712–15; disc. 715
Caruthers, R. S. Varistor Modulators for Carrier Systems. (54-218)..........................291–8
Caruthers, R. S.; K. E. Appert, W. S. Chaskin. Application and Transmission Features of a New 12-Channel Open-Wire Carrier System. (54-58)..18–27
Caywood, W. P.; D. Zeheb. A Symbolic Method for Synthesis of 2-Terminal Switching Circuits. (54-394)..690–3
Chaney, W. G.; C. R. Kraus, A. T. Steelman. Radio Transmission of Narrow-Band Mobile Radio Systems at 40-Megacycles. (54-240).........302–07
Charny, A. L. Application of Toll Dialing in Pennsylvania. (54-273)..........................400–05
Chaskin, W. S.; K. E. Appert, R. S. Caruthers. Application and Transmission Features of a New 12-Channel Open-Wire Carrier System. (54-58)..18–27
Clark, W. R.; W. G. Amey, F. H. Krantz, A. J. Williams, Jr. Stabilized Power Supplies for Instrument Applications. (54-295).....484–8; disc. 489
Coe, R. H.; F. F. McClatchie. One Approach to a Video SHF Relay System. (54-329).......323–9
Concordia, C. Disc..........................175
Cone, D. I.; F. W. Wright. Disc..............434
Cory, S. I. A New Portable Telegraph Transmission Measuring Set. (54-105)................59–62
Cruser, V. I. Equipment and Mechanical Features of the AN/TRC-24 Radio Set. (54-406)......544–7
Curtin, J. R.; C. M. Hathaway, W. L. Davis. A New 12-Element Automatic Oscillograph and Application on the Bonneville Power System. (54-293)..307–12
Cusack, F. H. Telegraph Terminal AN/FGC-29 Equipment Features. (54-402)...........536–40
Cushing, T. D.; Emil Hopner. Comparative Propagation Studies on 250 and 450 Megacycles. (55-241)..318–20

D

Dakin, T. W. Disc..........................153, 197
Dakin, T. W.; H. M. Philofsky, W. C. Divens. Effect of Electric Discharges on the Breakdown of Solid Insulation. (54-70)............155–61; disc. 162
Danser, J. W.; D. F. Hazen, G. S. Zillis. A Private Microwave Radio System for Power Company Use. (54-407)..492–8
Davis, B. E.; I. H. Swift. An Analogue Computer Technique Using Magnetic Amplifiers. (54-389)..........................635–9; disc. 639
Davis, W. L.; C. M. Hathaway, J. R. Curtin. A New 12-Element Automatic Oscillograph and Application on the Bonneville Power System. (54-293)..307–12
Deans, William. Disc..........................580
Decker, R. O. Alteration of the Dynamic Response of Magnetic Amplifiers by Feedback. (54-376)..........................658–63; disc. 664
Decker, R. O. Disc..........................225
Decker, R. O.; W. F. Horton. Disc..............430
Dillard, J. K.; C. J. Baldwin, Jr. Rectifier Arc-Back Study on the Analogue Computer. (54-43)..........................198–207; disc. 207
Divens, W. C.; T. W. Dakin, H. M. Philofsky. Effect of Electric Discharges on the Breakdown of Solid Insulation. (54-70)............155–61; disc. 162
Doigan, Paul; L. C. Whitman. Calculation of Life Characteristics of Insulation. (54-72)..........................193–7; disc. 198
Douglas, John F. H. Disc..........................233
Dunnegan, T., Jr.; J. D. Harnden, Jr. The Cyclic Integrator—A Device for Measuring the Frequency Response of Magnetic Ampliers. (54-260)..358–65
Dzimianski, J. W.; T. B. Jones. Characteristics of the High Current Argon Arc with Various Electrode Materials. (54-510).............665–9; disc. 670

E

Eckersley, Alfred. Disc..........................423
Ekstein, H.; T. L. Gilbert. Disc..............634
Ellert, F. J. Disc..........................616
Evans, W. G. Saturating Transformer Reference Circuit. (54-388)................610–16; disc. 616
Ewing, A. G.; F. W. Frazee, D. Welling. Mechanical Aspects and Component Features of a New 12-Channel Open-Wire Carrier System. (54-102)..75–81

F

Fahrnkopf, D. C.; W. E. Pakala. Disc. 424
Ferguson, P. R.; J. M. Moses. Disc. 502
Feth, G. C.; L. A. Finzi. Disc. 451
Finzi, L. A. Disc. 474, 639
Finzi, L. A.; G. C. Feth. Disc. 451
Finzi, L. A.; R. R. Jackson. The Operation of Magnetic Amplifiers with Various Types of Load:
  Part I. Load Currents for Given Angle cf Firing (54-237) 270–8
  Part II. Controlling the Angle of Firing. (54-238) 279–88
Fitchen, F. C.; F. P. Schwieg, W. H. Tucker. A New Standard for Direct Current. (54-354) .... 655–8; disc. 658
Frank, Max; S. Rabotnick, J. R. Walker. Analysis of a Single-Core Magnetic Amplifier with Real Rectifier and Core Functions. (54-158) .... 182–8; disc. 188
Fraser, W. Disc. 207
Frazee, F. W.; A. G. Ewing, D. Welling. Mechanical Aspects and Component Features of a New 12-Channel Open-Wire Carrier System. (54-102) 75–81
Frick, C. W. The Quasi-Peak Voltmeter. (54-204) .... 417–23; disc. 424

G

Gibson, J. E.; P. F. Ordung, B. J. Shinn. Closed-Loop Automatic Phase Control. (54-268) 375–81
Gilbert, T. L.; H. Ekstein. Disc. 634
Gluckman, H. P. Disc. 475

H

Hall, H. J. An Automatic Voltage Control System for Electrical Precipitators. (54-175) 124–7
Halverson, Gilbert. Electrical Properties of Microcrystalline Selenium. (54-163) 38–45
Hamilton, B. H. Some Applications of Semiconductor Devices in the Feedback Loop of Regulated Metallic Rectifiers. (54-418) 640–5
Hammond, S. B.; J. S. Carroll. The Measurement of Electrostatic Potential Attributable to Net Ion Space Charge in Air. (54-511) .... 712–15; disc. 715
Hansen, A., Jr.; F. R. Sias, J. R. Macintyre. A Tungsten Resistance Thermometer. (54-16) 66–9
Harnden, J. D., Jr.; T. Dunnegan, Jr. The Cyclic Integrator—A Device for Measuring the Frequency Response of Magnetic Amplifiers. (54-260) . 358–65
Hathaway, C. M.; W. L. Davis, J. R. Curtin. A New 12-Element Automatic Oscillograph and Application on the Bonneville Power System. (54-293) 307–12
Hawthorne, E. I. Flow of Energy in Synchronous Machines. (54-109) 1–10
Hays, J. B.; W. R. Bullard, H. O. Saunders. Electrical Protection of Telephone Systems. (54-289) 385–8
Hazen, D. F.; J. W. Danser, G. S. Zillis. A Private Microwave Radio System for Power Company Use. (54-407) 492–8
Hellerman, H. Some Transistor Building Blocks for Analogue Computers. (54-270) 410–13
Hermont, A. J. Precision Measurement of Phasors in Electric Networks as a Function of Frequency. (54-292) 365–75
Higgins, Thomas J.; W. Kneen. Basic Theory and Experimental Verification of the A-C Galvanometer. (54-171) .... 235–41; disc. 241
Higgins, Thomas J.; C. M. Siegel. Equations for Determining Current Distribution Among the Conductors of Busses Comprised of Double-Channel Conductors. (54-467) 489–93
Hoagland, Albert S. Magnetic Drum Recording of Digital Data. (54-271) 381–5
Hoffmann, J. P. New Military Carrier Telephone Systems Equipment Features. (54-410) 509–15
Hollander, Gerhard L. A Digital Data-Recorder for Dense Storage of Continuous Voltages. (54-318) 253–9
Hollander, Gerhard L. Bibliography on Data Storage and Recording. (54-169) 49–58
Hopner, Emil; T. D. Cushing. Comparative Propagation Studies on 250 and 450 Megacycles. (54-241) 318–20
Horton, W. F.; R. O. Decker. Disc. 430
Hubbard, R. M. An Operational-Type Magnetic Amplifier for Airborne Servo-Control Systems. (54-247) .... 425–30; disc. 430
Huber, G. H.; W. F. Miller, C. W. Schramm. New Military Carrier Telephone Systems. (54-386) 515–25
Hughes, Gerald E.; H. A. Miller. Fast-Response Magnetic Amplifiers. (54-159) 69–75
Huhta, Hoobert. Flux Resetting Characteristics of Several Magnetic Materials. (54-161) 111–14
Humphrey, T. R. A Radio Relay Remote-Control System for Frequency-Modulation Broadcast Stations. (54-167) 128–31
Huntley, H. R. Disc. 208

J

Jackson, R. R. Disc. 616
Jackson, R. R.; L. A. Finzi. The Operation of Magnetic Amplifiers with Various Types of Loads:
  Part I. Load Currents for Given Angle cf Firing. (54-237) 270–8
  Part II. Controlling the Angle of Firing. (54-238) 279–88
Jerrard, R. P.; H. Poritsky. Eddy-Current Losses in a Semi-Infinite Solid due to a Nearby Alternating Current. (54-106) 97–106
Johannessen, P. R. Analysis of Magnetic Amplifiers by the Use of Difference Equations. (54-509) .... 700–11; disc. 711
Johnson, L. J.; S. E. Rauch. Magnetic Frequency Multipliers. (54-248) .... 448–51; disc. 451
Johnston, J. H.; R. S. Boykin, S. D. Bedrosian. Considerations for Development of New Military Carrier Telephone Systems. (54-408) 503–06
Johnston, Leith; R. C. Stiles. A Fully Automatic Teletypewriter Switching Center for Military Use. (54-122) 27–37
Jones, Byron M. Disc. 488
Jones, T. B.; J. W. Dzimianski. Characteristics of the High Current Argon Arc with Various Electrode Materials. (54-510) .... 665–9; disc. 670
Jury, Eliahu I. Analysis and Synthesis of Sampled-Data Control Systems. (54-298) 332–46

K

Kaplan, Henry; G, Wolff. Techniques for Measuring Cascaded Self-Saturating Magnetic-Amplifier Performance. (54–325) .... 581—4
Karnaugh, M. Disc. 146
Kelly, H. P. Differential Phase and Gain Measurements in Color Television Systems. (54-334) 565–9
Keonjian, Edward; J. S. Schaffner. Shaping of the Characteristics of Temperature-Sensitive Elements (54-269) 396–400
Kingsley, H. F. X. A New Cable Design for Military Carrier Telephone Systems. (54-409) 506–09
Kirk, Don; K. A. Simons, H. J. Arbeiter. Line Amplifiers for Community Television Systems. (54-101) .... 214–19; disc. 220
Kneen, William; T. J. Higgins. Basic Theory and Experimental Verification of the A-C Galvanometer. (54-171) .... 235–41; disc. 241
Koenig, H. E. Application of Network Theory to the Analysis of Rotating Machinery:
  Part I. Synchronous and Asynchronous Machines. (54-87) 162–9
  Part II. Commutating Machines. (54-88) .. 169–74
  Discussions 174
Krantz, F. H.; W. G. Amey, W. R. Clark, A. J. Williams, Jr. Stabilized Power Supplies for Instrument Applications. (54-295) .. 484–8; disc. 489
Kraus, C. R.; W. G. Chaney, A. T. Steelman. Radio Transmission of Narrow-Band Mobile Radio Systems at 40-Megacycles. (54-240) 302–07
Kron, Gabriel. Electric Circuit Models of the Nuclear Reactor. (54-199) 259–65
Ku, Y. H. Acceleration Plane Method for Analysis of a Circuit with Nonlinear Inductance and Nonlinear Capacitance. (54-514) 619–26
Ku, Y. H. Analysis of Nonlinear Coupled Circuits. (54-515) .... 626–30; disc. 630

L

Langford, R. C. Disc. 416
Larson, E. L. C.; G. M. L. Sommerman, C. J. Bute. Impulse Ionization and Breakdown in Liquid Dielectrics. (54-69) .... 147–53; disc. 154
Laube, O. T. Technical Considerations Relating to Marine Radiotelephone Communication for Safety (54-242) 347–53
LeBel, C. J. Disc. 655
Lee, C. Y. Switching Functions on an n-Dimensional Cube. (54-207) 289–91
Levine, R. H.; A. Mack. A New Multichannel Teletype for Use on Long-Range High-Frequency Radio Systems. (54-401) 540–4
Lewis, Laurel J. Harmonic Analysis for Nonlinear Characteristics. (54-512) .... 693–9; disc. 699
Li, Kam; H. P. Schwan. Measurements of Materials with High Dielectric Constant and Conductivity at Ultrahigh Frequencies. (54-524) 603–07
Liao, T. W. Disc. 161, 715
Liao, T. W.; J. G. Anderson. Disc. 153
Lisser, S.; D. W. C. Shen. An Analogue Computer for Automatic Determination of System Swing Curves. (54-275) .... 475–83; disc. 483
Lord, Harold W. Magnetic-Amplifier Circuits with Full-Wave Output and Half-Wave Control Signals. (54-249) 265–70
Lufcy, C. W.; H. H. Woodson. Design Considerations of the Half-Wave Bridge Magnetic Amplifier. (54-182) .... 220–5; disc. 225

M

Macintyre, J. R.; F. R. Sias, A. Hansen, Jr. A Tungsten Resistance Thermometer. (54-16) 66–9
Mack, Alfred; R. H. Levine. A New Multichannel Teletype for Use on Long-Range High-Frequency Radio Systems. (54-401) 540–4
Mackenzie, W. F. Impedance and Induced Voltage Measurements on Iron Conductors. (54-469) .... 577–80; disc. 580
Mark, M. Torque Requirements of a Radar Antenna. (54-6) 15–18
McAninch, O. G. Thermocouple-Type Ammeters for Use at Very High Frequencies. (54-170) .... 241–5; disc. 245
McClatchie, F. F.; R. H. Coe. One Approach to a Video SHF Relay System. (54-239) 323–9
McConnell, H. M. Eddy-Current Phenomena in Ferromagnetic Materials. (54-146) .. 226–33; disc. 234
McCutchen, James M. Disc. 424
McDonough, T. J.; W. T. Smith. Some Engineering Considerations in the Design of Telephone Systems to Serve Predominantly Rural Areas. (54-103) 208–13
McFerran, J. B.; R. E. Morgan. Pulse Relaxation Amplifier—A Low-Level D-C Magnetic Amplifier. (54-198) 245–9
McMaster, Robert C. Disc. 669
Meahl, H. R. Disc. 245
Menyuk, N.; D. R. Brown, D. A. Buck. A Comparison of Metals and Ferrites for High-Speed Pulse Operation. (54-361) .... 631–4; disc. 634
Merrill, J. L., Jr. Theory of E-Type Repeaters. (54-286) 443–7
Miller, Harvey A.; G. E. Hughes. Fast-Response Magnetic Amplifiers. (54-159) 69–75
Miller, J. H. A Simplified Standard Cell Comparator. (54-208) .... 413–15; disc. 416
Miller, J. H. Disc. 651
Miller, W. F.; G. H. Huber, C. W. Schramm. New Military Carrier Telephone Systems. (54-386) 515–25
Molnar, Imre. Some Recent Advances in the Economy of Routing Calls in Nationwide Toll Dialing. (54-503) 670–9
Morgan, R. E. The Biased Rectifier Amplifier—A Pulse Magnetic Amplifier. (54-364) 584–90
Morgan, R. E.; J. B. McFerran. Pulse Relaxation Amplifier—A Low-Level D-C Magnetic Amplifier. (54-198) 245–9
Moses, J. M.; P. R. Ferguson. Disc. 502

N

Narbut, Paul. Disc. 181
Nexon, Victor J. Industry Co-ordination of Microwave Communications Systems. (54-188) 250–3
Nolke, Frederick H. Interoffice Trunking and Signaling. (54-502) 590–7
Nordahl, John G. A New Ultrahigh-Frequency Multichannel Military Radio Relay System. (54-405) 526–31

O

O'Connor, R. R. Disc. 435
Oler, Charles B. The Use of Sphere Gaps at Radio Frequencies. (54-294) 329–32
O'Neil, S. J. Networks for Digital-to-Analogue Shaft-Position Transducers. (54-272) 456–66
Ordung, P. F.; J. E. Gibson, B. J. Shinn. Closed-Loop Automatic Phase Control. (54-268) 375–81
Orlando, V. A. A Transducer Using a Short-Circuit Rotor. (54-172) 131–4

P

Pakala, W. E.; D. C. Fahrnkopf. Disc. 424
Penney, G. W.; E. A. Sack, E. R. Wingrove. Frequency Response of a Resonant Dielectric Amplifier. (54-160) 119–24

Philofsky, H. M.; T. W. Dakin, W. C. Divens. Effect of Electric Discharges on the Breakdown of Solid Insulation. (54-70)............155–61; disc. 162
Pipes, Louis A. Analysis of Linear Time-Varying Circuits by the Brillouin-Wentzel-Kramers Method. (54-183)..............................93–6
Pipes, Louis A. Forced Oscillations of Nonlinear Circuits. (54-297)..........................352–8
Pittman, G. F., Jr. Instantaneous Electrical Characteristics of Selenium Rectifiers. (54-164)........45–9
Poritsky, H.; R. P. Jerrard. Eddy-Current Losses in a Semi-Infinite Solid due to a Nearby Alternating Current. (54-106).......................97–106

Q

Queen, J. L.; W. A. Yates. Sheet and Plated-Metal Measurements with a Phase-Angle-Type Probe. (54-173)..............................138–42
Quinlan, A. L. Automatic Percussion Welding. (54-347)..............................561–5

R

Rabotnick, Saul; M. Frank, J. R. Walker. Analysis of a Single-Core Magnetic Amplifier with Real Rectifier and Core Functions. (54-158).182–8; disc. 188
Radus, R. J. Single-Ended Saturable Reactor Circuit with Quiescent Current Compensation. (54-393)................................597–603
Rauch, S. E.; L. J. Johnson. Magnetic Frequency Multipliers. (54-248)..........448–51; disc. 451
Riaz, M. Disc...............................663
Ribe, M. L.; S. P. Brown. Considerations for a New Military Radio Relay System. (54-404)...547–52
Rinker, L. E. Disc.......................580, 658
Roberts, R. W. Magnetic Characteristics Pertinent to the Operation of Cores in Self-Saturating Magnetic Amplifiers. (54-392)...........682–9; disc. 690
Rorden, H. L. Disc.............................424
Rose, Arthur F. Negative-Impedance Telephone Repeaters—Application in the Bell System. (54-222)..............................430–4; disc. 434
Rosenstein, Allen B. Precision High-Current Computer Power Supplies. (54-313)..............405–09

S

Sack, E. A.; G. W. Penney, E. R. Wingrove. Frequency Response of a Resonant Dielectric Amplifier. (54-160)..............................119–24
Saunders, H. O.; W. R. Bullard, J. B. Hays. Electrical Protection of Telephone Systems. (54-289)..................................385–8
Saunders, R. M. Digital Computers as an Aid in Electric-Machine Design. (54-168)..........................189–91; disc. 191
Saunders, Robert M. Disc............174, 456, 630
Sayer, W. H. An Ultrahigh-Frequency Transmitter Employing Klystron Power Amplifiers. (54-166)..................................114–18
Schaefer, David H. A Rectifier Algebra. (54-516)..................................679–82
Schaffner, J. S.; E. Keonjian. Shaping of the Characteristics of Temperature-Sensitive Elements. (54-269).........................396–400
Scheideler, A. L. Significant Measurements for Determining the Stability of High-Temperature Magnet-Wire Insulation. (54-71)......177–81; disc. 181
Schmidt, A., Jr. Disc..........................207
Schnitzler, A. D. Derivative-Controlled Magnetic Amplifiers. (54-375)....................569–76
Schramm, C. W.; G. H. Huber, W. F. Miller. New Military Carrier Telephone Systems. (54-386)..................................515–25
Schultheiss, C. E. High Speed Teletypewriter Equipment for the Armed Services. (54-114).....88–93
Schwan, H. P.; K. Li. Measurements of Materials with High Dielectric Constant and Conductivity at Ultrahigh Frequencies. (54-524).........603–07
Schwieg, F. P.; F. C. Fitchen, W. H. Tucker. A New Standard for Direct Current. (54-354)..........................655–8; disc. 658
Shen, D. W. C.; S. Lisser. An Analogue Computer for Automatic Determination of System Swing Curves. (54-275)....................475–83; disc. 483
Shinn, B. J.; P. F. Ordung, J. E. Gibson. Closed-Loop Automatic Phase Control. (54-268).......375–81
Shullman, J. M. Accurate Tachometry Methods with Electronic Counters. (54-290)....452–5; disc. 456
Sias, F. R.; J. R. Macintyre, A. Hansen, Jr. A Tungsten Resistance Thermometer. (54-16)......66–9
Siegel, Clifford M.; T. J. Higgins. Equations for Determining Current Distribution Among the Conductors of Busses Comprised of Double-Channel Conductors. (54-467).....................489–92
Silsbee, F. B. Disc............................416
Simons, K. A.; D. Kirk, H. J. Arbeiter. Line Amplifiers for Community Television Systems. (54-101)..........................314–19; disc. 220
Smethurst, J. O. E-Type Telephone Repeaters—Description, Equipment, and Testing. (54-287)..................................435–43
Smith, D. H. The Suitability of the Silicon Alloy Junction Diode as a Reference Standard in Regulated Metallic Rectifier Circuits. (54-417)..........................645–51; disc. 651
Smith, Howard G. High-Gain Side-Firing Helical Antennas for Ultrahigh-Frequency Television Broadcasting. (54-165)..................135–8
Smith, J. J. The Use of Instantaneous Point Sources or Green's Functions in Evaluating Electromagnetic Fields. (54-162)........................82–8
Smith, Lester C. Disc..........................214
Smith, Warner T.; T. J. McDonough. Some Engineering Considerations in the Design of Telephone Systems to Serve Predominantly Rural Areas. (54-103)............................208–13
Sommerman, G. M. L. Disc...................161
Sommerman, G. M. L.; C. J. Bute, E. L. C. Larson. Impulse Ionization and Breakdown in Liquid Dielectrics. (54-69)..............147–53; disc. 154
Steelman, A. T.; C. R. Kraus, W. G. Chaney. Radio Transmission of Narrow-Band Mobile Radio Systems at 40 Megacycles. (54-240).........302–07
Stewart, Chandler. Measurement of the Quality Factor of Inductor Cores. (54-291)............298–302
Stiles, R. C.; L. Johnston, A Fully Automatic Teletypewriter Switching Center for Military Use. (54-122)................................27–37
Storch, Leo. The Transmission Matrix of N Alike Cascaded Networks. (54-513)................616–8
Storm, H. F. Magnetic Amplifiers with Inductive D-C Load. (54-316)..............466–74; disc. 474
Stout, Thomas M. Disc........................699
Swift, I. H.; B. E. Davis. An Analogue Computer Technique Using Magnetic Amplifiers. (54-389)...........................635–9; disc. 639

T

Tucker, W. H.; F. C. Fitchen, F. P. Schwieg. A New Standard for Direct Current. (54-354)..........................655–8; disc. 658

V

Veinott, C. G. Disc...........................174

W

Walker, J. R.; M. Frank, S. Rabotnick. Analysis of a Single-Core Magnetic Amplifier with Real Rectifier and Core Functions. (54–158)..182–8; disc. 188
Washburn, S. H. Disc.........................146
Watt, Gordon J. Disc.........................233
Wedel, W. H.; W. S. Ames. Type-N1 Carrier on Radio and Coaxial Cable. (54-214).............312–18
Welling, Dale; A. G. Ewing, F. W. Frazee. Mechanical Aspects and Component Features of a New 12-Channel Open-Wire Carrier System. (54-102).....................................75–81
Whitman, L. C.; P. Doigan. Calculation of Life Characteristics of Insulation. (54-72)..193–7; disc. 198
Wildes, Karl L. Disc..........................175
Williams, A. J., Jr; W. G. Amey, F. H. Krantz, W. R. Clark. Stabilized Power Supplies for Instrument Applications. (54-295)..........484–8; disc. 489
Windeler, A. S. Polyethylene-Insulated Telephone Cable. (54-104)........................106–11
Wingrove, E. R.; G. W. Penney, E. A. Sack. Frequency Response of a Resonant Dielectric Amplifier. (54-160).............................119–24
Winograd, H. Disc............................207
Wolf, E. Mark. Disc..........................509
Wolff, Gerald; H. Kaplan. Techniques for Measuring Cascaded Self-Saturating Magnetic-Amplifier Performance. (54-325).....................581–4
Woods, W. W. Thickness Gauge for Dielectric Materials. (54-296)........................320–3
Woodson, H. H. Disc..........................711
Woodson, H. H.; C. W. Lufcy. Design Considerations of the Half-Wave Bridge Magnetic Amplifier. (54-182)..........................220–5; disc. 225
Works, C. N. Disc............................153
Wright, F. W.; D. I. Cone. Disc.................434

Y

Yates, W. A.; J. L. Queen. Sheet and Plated-Metal Measurements with a Phase-Angle-Type Probe. (54-173)..............................138–42
Young, K. D.; L. B. Bogan. Simplified Transmission Engineering in Exchange Cable Plant Design. (54-411).................................498–502

Z

Zeheb, D.; W. P. Caywood, A Symbolic Method for Synthesis of 2-Terminal Switching Circuits. (54-394)..................................690–3
Zenner, W. J.; A. S. Benjamin. A Step Forward in Printing Telegraphy. (54-115)............10–15
Zillis, G. S.; D. F. Hazen, J. W. Danser. A Private Microwave Radio System for Power Company Use. (54-407)...............................492–8

Flow of Energy in Synchronous Machines

E. I. HAWTHORNE
MEMBER AIEE

Synopsis: The flow of energy is traced in a typical, somewhat idealized 3-phase alternator, using an adaptation of Poynting's vector together with a discussion of this procedure. Following the specification of the electromagnetic field, quantitative analysis of the flow of energy in the gap, at the surface of the rotor, and into and out of the stator conductors is presented under no load, at a unity-power-factor load and then for any power-factor load. Qualitative description of the flow of energy in the intermediate regions and to the output terminals of the machine is then given, together with a brief discussion of the extension of this analysis to other machines and to less idealized situations.

A STUDY of the flow of energy in a synchronous polyphase a-c machine is carried out on the basis of the analysis of the electromagnetic field within the machine. This constitutes the second phase of the study of energy flow in rotating machinery, following a previous analysis of the d-c machine.[1] The objectives of this investigation are the same as those stated in the above paper: to afford a better understanding of the operation of machines from an educational and analytical standpoint, and possibly to be of some value in the improvement of their design.

The basic assumptions and the resulting limitations of the analysis remain essentially the same as in the d-c case. The major weakness is the postulated linearity of the problem, which neglects saturation and hysteresis effects. Since the inclusion of ohmic losses complicates the analysis somewhat without essentially altering the basic principles involved, the conductors, as well as the iron, will be assumed to have negligible losses. (Conductor losses could be easily included in a manner analogous to their treatment in the d-c machine in reference 1.) Steady-state operation at constant speed of a cylindrical rotor machine with a uniform, small gap will be considered. A uniform symmetric polyphase stator winding will be assumed, composed of many closely spaced conductors, and effects of nonuniformity produced by the presence of slots and ventilating ducts will be ignored. The permeability of iron will be assumed much greater than that of air. A 2-pole d-c excitation will be assumed, whose magnetic field intensity in the gap has a sinusoidal space variation. The flow of energy associated with the maintenance of the d-c excitation will again be ignored. (This corresponds to permanent magnet excitation and was similarly done in the analysis of the d-c machine in reference 1. The flow of energy associated with the d-c excitation of the rotor is analogous to the flow of energy in a simple reactor, and contributes a negligible amount to the over-all picture of energy flow in an a-c machine.)

In particular, a 3-phase wye-connected full-pitch single-layer alternator with two conductors per turn will be analyzed at no load, under a unity-power-factor load, and under any power-factor load. Extension of the analysis to any other synchronous machine will then be discussed.

The main new aspect of the problem, in contrast to the d-c machine, is that time-varying fields are present. Consequently, additional components of the energy flow vector[2] appear and must be included in the analysis. However, because of the relatively slow time variation of the field vectors at power frequencies, certain of these components may be safely neglected, as discussed in Appendix I, in line with the arguments that underlie the concepts of energy flow; that is, with a given postulate of electromagnetic energy, all flow vectors having the same, or nearly the same, required value of divergence are equally valid and equally significant physically.[2]

The approach will again consist of first specifying the electromagnetic fields present in the region of the machine under consideration. Employing the conventional postulate of electromagnetic energy a modified suitable form of Poynting's vector[2] will then be constructed from the

Nomenclature

The rationalized absolute meter-kilogram-second-MKSA) system of units is presumed. Vector quantities have the usual bold-face-type designation. Co-ordinate direction of vector components is designated by superscripts.

a = subscript designating the stator as source
A = phase belt designation
b = subscript designating the rotor as source
B = phase belt designation
$\mathbf{B}$ = magnetic flux density
c = subscript designating active (cosine) components
C = phase belt designation
d = designation of a differential, or derivative
D = diameter of the rotor
$\mathbf{E}$ = electric field intensity
f = frequency
F = functional designation
g = gap width
G = normalized flow vector function (see equations 23 and 24)
h = conductor height
$\mathbf{H}$ = magnetic field intensity
i = instantaneous current
I = rms value of current
$\mathbf{J}$ = current density
k = an integer
k_d = distribution factor of the stator winding
m = integral designation of phase belts
n = angular speed, revolutions per minute
N = number of turns
p = power
p_0 = power output of the machine
$\mathbf{P}$ = energy flow vector
q = vertical co-ordinate in conductors
r = radial co-ordinate
R = normalized flow vector function (see equations 36 and 37)
s = subscript designating reactive (sine) components
$\mathbf{s}$ = distance
S = normalized flow vector function (see equations 28 and 29)
t = time
T = normalized flow vector function (see equations 47, 48, 54, and 55)
u = normalized axial co-ordinate z (see equation 44)
U = normalized flow vector function (see equations 71, 72, 74, and 75)
$\mathbf{v}$ = velocity of the surface of the rotor
V = rms value of generated voltage per phase
x = reference co-ordinate
y = reference co-ordinate
z = axial co-ordinate
Z = number of stator conductors
α = ratio of gap width to rotor radius
β = power-factor angle of lag
δ = angular displacement between successive conductors (see equation 44)
$\epsilon(\epsilon_0)$ = permittivity (of free space)
θ = angular co-ordinate
λ = axial length of the machine
$\mu(\mu_0)$ = permeability (of free space)
π = ratio of circumference to radius = 3.14159 . . .
ρ = electric charge density
τ = normalized time co-ordinate (see equation 2)
ω = angular speed of the rotor
∂ = designation of a partial derivative

Paper 54-109, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted September 8, 1953; made available for printing November 20, 1953.

E. I. HAWTHORNE is with the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pa.

The author gratefully acknowledges the inspiring guidance and assistance of Dr. J. G. Brainerd in the study of the problem and the critical review of the original work by Dr. Y. H. Ku, both of the University of Pennsylvania.

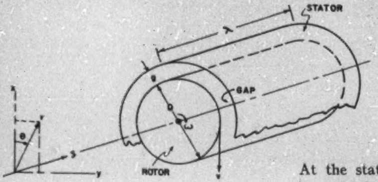


Fig. 1. Co-ordinate system

interaction of the electric and magnetic field components present, omitting, in line with the discussion in Appendix I, those components which by virtue of having zero or negligibly small divergence contribute little to the predominant mode of energy flow in the machine.

Alternator at No Load

The revolving steady magnetic field of the rotor appears as a sinusoidal, time-varying quantity with respect to the stationary co-ordinate system of the stator, shown on Fig. 1. This magnetic field intensity in the gap, radially directed for the assumed high-permeability iron, may be written as

$$H_b^r(\theta,\tau)=H'\cos(\theta-\tau) \qquad (1)$$

for a suitably chosen phase reference. H' is the amplitude of the wave and τ is the normalized time co-ordinate

$$\tau=\omega t=(2v/D)t \qquad (2)$$

where the angular frequency ω is

$$\omega=2v/D=2\pi f=2\pi n/60 \qquad (3)$$

For a complete period, $t=\pi D/v$, corresponding to $\tau=2\pi$.

As in the case of the d-c machine,[1] a charge distribution exists on the stator conductors, setting up an electric field which just counterbalances the $\mathbf{v}\times\mathbf{B}$ force exerted by the revolving field on the conductor charges. Alternately, choosing the viewpoint of the "flux linking" induction law, rather than that of the "flux cutting" $\mathbf{v}\times\mathbf{B}$ law, a total $\overline{\mathbf{E}}_a$ field may be said to be associated with the stator conductors, composed of two components defined as follows

$$\mathbf{E}_a=\mathbf{E}_1+\mathbf{E}_2 \qquad (4)$$

$$\text{div }\mathbf{E}_1=-\frac{\rho}{\epsilon_0};\ \text{curl }\mathbf{E}_1=0;\ \text{div }\mathbf{E}_2=0;$$

$$\text{curl }\mathbf{E}_2=-\mu_0\frac{\partial\mathbf{H}}{\partial t} \qquad (5)$$

The $\mathbf{E}_2$ component is readily identified with the $\mathbf{v}\times\mathbf{B}$ effect and may be calculated from it in the gap as

$$E_2^r=0;\ E_2^\theta=0;\ E_2^s=\mu_0 vH'\cos(\theta-\tau) \qquad (6)$$

At the stator conductors, where $\mathbf{E}_a$ is zero, the $\mathbf{E}_1$ and $\overline{\mathbf{E}}_2$ components cancel.

Strictly speaking H' and v vary with r across the gap space. For the present it will suffice to consider a small gap $g << D$. Variation with radial distance is treated in Appendix II, where it is shown to be of minor consequence. It is also shown there, that even for an appreciable width of gap, E_2^s is independent of r. We have then

$$E_2^s(\theta,\tau)=(\mu_0\omega DH'/2)\cos(\theta-\tau) \equiv E\cos(\theta-\tau) \qquad (7)$$

where E is a constant amplitude.

The energy flow vector

$$\mathbf{P}=\mathbf{E}\times\mathbf{H} \qquad (8)$$

is thus seen to exist at no load in the gap, where it has two parts, $\mathbf{E}_1\times\mathbf{H}_b$ and $\mathbf{E}_2\times\mathbf{H}_b$. The former vector, by arguments presented in Appendix I, has a negligibly small divergence in the gap, and therefore represents a merely circulatory flow of energy, whose conceptual origin lies in the basic formulation of the postulate of electromagnetic energy flow.[2] This vector is hence of no consequence in the study of the overall transfer of power, and may be disregarded. The latter part is a θ-directed vector of magnitude

$$P_{2b}^\theta(\theta,\tau)=[\mu_0\omega D(H')^2/4]\times[1+\cos(2\theta-2\tau)] \qquad (9)$$

Since H_b^r and E_2^s reverse polarity simultaneously, this vector has the same θ sense everywhere. Its divergence is seen to be $-\mu_0\omega(H')^2\sin(2\theta-2\tau)$, and has a zero average value. The flow of energy associated with it is circulatory in nature and no net transfer of power takes place. With respect to the stator, the flow vector of equation 9 has two components: one constant, and one a wave traveling in the $+\theta$ direction with a velocity v. With respect to the rotor the situation is reversed, the latter being a constant component, and the former a wave traveling in the $-\theta$ direction. From either standpoint, the analysis of the no-load condition yields the expected conclusion of zero net power transfer, and a merely circulatory flow of energy.

The conclusions of this section hold whether or not there are any stator conductors. Their presence simply serves to "generate" the additional $\mathbf{E}_1$ field, which accounts for the developed open-circuit terminal voltage of the machine.

Unity-Power-Factor Load—General Discussion

The electromagnetic field and the flow of energy in the machine are modified considerably under load and depend on the power factor, that is, the phase relation between the rotor field $\mathbf{H}_b$ and the field $\mathbf{H}_a$ of the stator currents. The special case where the peak of the $\mathbf{H}_b$ field passes through the center of the stator phase belt at the instant when the current in that belt is a maximum, corresponding closely to a unity-power-factor load, will first be treated.

The sinusoidal stator currents modify the field in the machine primarily by contributing a magnetic field intensity $\mathbf{H}_a$. The contribution of an electric field $\mathbf{E}_a$ will be neglected. The path of the stator magnetomotive force, especially at unity power factor, is largely in air, and the stator magnetomotive force is usually considerably smaller than the rotor magnetomotive force. (At worst, this analysis may be said to provide a first approximation to the flow of energy in a lightly loaded machine.) Inclusion of the effect of $\mathbf{E}_a$ in the analysis is straightforward, as was demonstrated in the case of the d-c machine,[1] but it complicates the equations somewhat without appreciably affecting the resultant conclusions with respect to energy flow.

The interaction of the $\mathbf{H}_b$ field with the $\mathbf{E}$ field still yields a circulatory flow of energy, as evident from the discussion of the unloaded machine. (This is true even if $\mathbf{E}_a$ is appreciable, since the argument with respect to $\mathbf{E}_1$ is unchanged, and the resultant modification of the amplitude and phase of $\mathbf{E}_2$ does not alter the conclusions of the previous section.) The flow-vector components to be studied, that appear under load, are then

$$\mathbf{P}_1=\mathbf{E}_1\times\mathbf{H}_a;\ \mathbf{P}_2=\mathbf{E}_2\times\mathbf{H}_a \qquad (10)$$

The vector $\mathbf{P}_1$ describes the energy that is "generated" within the stator conductors and flows to the machine output terminals. The vector $\mathbf{P}_2$ describes the flow of energy from the rotor structure, where mechanical energy is being supplied, across the gap, and to the stator conductors, where it supplies the $\mathbf{P}_1$ energy, equal and opposite to it there. It will be observed that at the rotor surface the $\mathbf{P}_2$ component, representing an interaction of the stator magnetomotive force with the electric field induced by the rotor magnetomotive force, is directly related to the mechanical torque present. Space does not permit a further discussion of the mechanical aspects of energy flow and conversion.

The Magnetic Field of Stator Currents in the Gap

The magnetic field intensity $\mathbf{H}_a$ in the gap will first be specified with the aid of Fig. 2, where the developed stator phase belts are shown for an assumed *ABC* sequence. The time reference is suitably chosen, for the assumed unity-power-factor condition, so that the magnetic field intensity H_b^r of the rotor, drawn for $\tau=0$, is at its positive peak in the center of the belt *A*, in whose conductors the currents (into the paper) are maximum at this instant. Thus

$$i_A=I\sqrt{2}\cos\tau \qquad (11)$$

$$i_B=I\sqrt{2}\cos(\tau-2\pi/3) \qquad (12)$$

$$i_C=I\sqrt{2}\cos(\tau-4\pi/3) \qquad (13)$$

The field intensity $\mathbf{H}_a(\theta,\tau)$ in the gap, predominantly radial, is composed of three parts

$$H_a^r(\theta,\tau)=H_A^r(\theta,\tau)+H_B^r(\theta,\tau)+H_C^r(\theta,\tau) \qquad (14)$$

due to currents in phase belts *A—A′*, *B—B′* and *C—C′*, respectively. These three components have the same relative dependence on θ but differ in space position and in time phase. The H_A^r component is shown on Fig. 2 at $\tau=0$ as a function of θ. It is called $F_A(\theta)$, and can be readily described mathematically for the assumed large number of conductors. In general

$$H_A^r(\theta,\tau)=F_A(\theta)\cdot\cos\tau \qquad (15)$$

$$\begin{aligned}H_B^r(\theta,\tau)&=F_B(\theta)\cdot\cos(\tau-2\pi/3)\\&=F_A(\theta-2\pi/3)\cdot\cos(\tau-2\pi/3)\end{aligned} \qquad (16)$$

$$\begin{aligned}H_C^r(\theta,\tau)&=F_C(\theta)\cdot\cos(\tau-4\pi/3)\\&=F_A(\theta-4\pi/3)\cdot\cos(\tau-4\pi/3)\end{aligned} \qquad (17)$$

The maximum value *H* of H_A^r may readily be shown to be

$$H=(\sqrt{2}ZI)/(12g) \qquad (18)$$

The complete field H_b^r may now be calculated. Its expression for the six regions specified by the integer *m* (shown on Fig. 2), where

$$m=0,1,2,3,4,5 \text{ for belts } A, C', B, A', C, B', \text{ respectively} \qquad (19)$$

is

$$H_a^r(\theta,\tau,m)=H\sqrt{3}\sin(\tau-m\pi/3)-\frac{6H}{\pi}(\theta-m\pi/3)\cos(\tau-m\pi/3) \qquad (20)$$

The above expression is more exact than the usual textbook approximation of the travelling stator field wave, which uses the fundamentals of the Fourier expansions of the $F(\theta)$'s only, and which is not sufficient to study the flow of energy properly. Thus the H_b^r field in the gap consists of two components, one independent of θ and the other varying linearly with θ for an assumed large number of stator conductors.

Flow of Energy in the Gap

The flow vector $\mathbf{P}_2$ of equation 10 is θ directed in the central region of the gap, but acquires a radial component in the vicinity of the conductors due to the curvature of the $\mathbf{H}_a$ lines. This P_2^θ component in the gap may be evaluated from the expressions for E_2^z and H_a^r of equations 7 and 20, respectively. Since the generated voltage per phase has an rms value of

$$V=(1/\sqrt{2})2E\lambda Zk_d/6 \qquad (21)$$

where k_d is the distribution factor, here equal to $3/\pi$ (see, for example, reference 3), we have, using equation 18

$$P_2^\theta=\frac{3VI}{\lambda_g}[G(\theta,m)+G'(\theta,\tau,m)] \qquad (22)$$

$$G(\theta,m)=\frac{\pi}{12\sqrt{3}}\sin(\theta-m\pi/3)-\frac{1}{6}(\theta-m\pi/3)\cos(\theta-m\pi/3) \qquad (23)$$

$$G'(\theta,\tau,m)=-\frac{\pi}{12\sqrt{3}}\sin[(\theta-m\pi/3)-2(\tau-m\pi/3)]-\frac{1}{6}(\theta-m\pi/3)\cos[(\theta-m\pi/3)-2(\tau-m\pi/3)] \qquad (24)$$

The P_2^θ flow vector consists of two parts whose normalized values are the *G* and *G′* functions above, where *G* is independent of time and represents the average value of P_2^θ at any θ. Note that the ratio $3VI/\lambda g$, which served to nor-

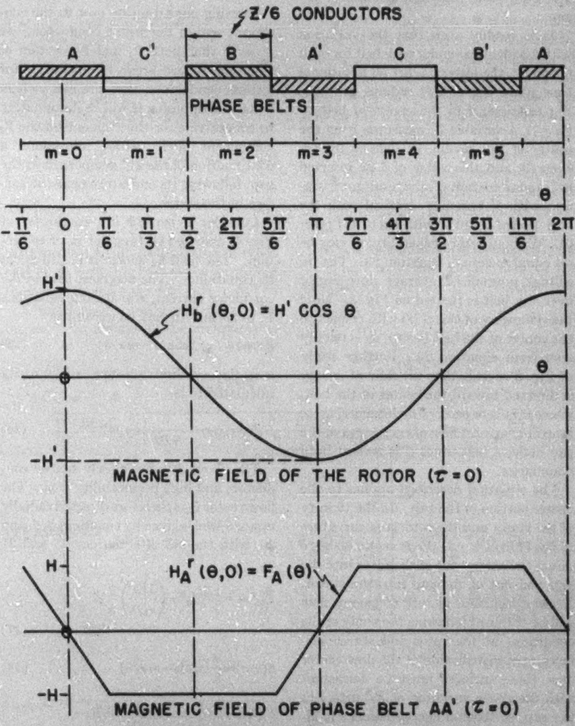


Fig. 2. Phase belts and magnetic field in the gap. Unity-power-factor condition

E_2^z field is into the paper

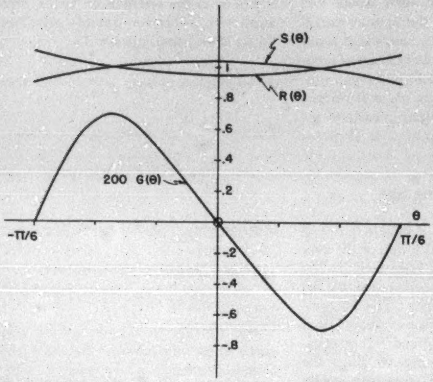


Fig. 3 (left). Normalized flow-vector functions, G (θ), R (θ), and S (θ). G (θ) is magnified 200 times

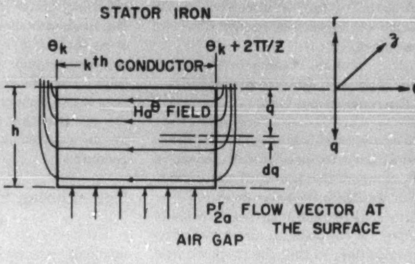


Fig. 4. A typical stator conductor

malize P_2^θ in equation 22, stands for the total power delivered by the machine divided by the cross-sectional area of the gap.

It is readily seen that the variation with θ is identical under each belt for both G and G', the latter having an additional time phase lag of $\pi/3$ radians per belt. For example, if in the region of belt C' ($m=1$), a variable θ', measured from the center of this region, be defined by $\theta=\theta'+\pi/3$, and this value of θ be inserted in G and G' for belt C', the resultant variation with θ' becomes identical with the variation of G and G' with θ in belt A ($m=0$). Without loss of generality, m may be set equal to zero in equation 23. The resultant normalized average flow vector over any belt is plotted on Fig. 3. Note the symmetry of this $G(\theta)$ with respect to the center of the belt ($\theta=0$), as is readily seen from equation 23. Further study of Fig. 3 reveals that the flow of energy is directed toward the center of the belt, where it falls to zero. Furthermore, there is no net outward flow of energy out of the gap under a belt, since G is zero at both boundaries.

The situation described applies to the center portion of the gap. In the vicinity of the stator and the rotor, the curvature of the **H** lines bends the **P** vector so as to direct the energy out of the rotor into the gap and out of the gap into the stator. A more complete picture of energy flow will be obtained following the study of the conditions at the rotor and stator surfaces; the distribution of the flow vector over these surfaces must be consistent with the above variation of P_2^θ with θ in the gap, so that an energy balance is always maintained, both on an over-all, integrated basis, and for each infinitesimal strip of the gap of width $Dd\theta/2$.

Energy Flow in the Region of the Stator Conductors

Turning our attention now to the conditions within the stator conductors, we observe that both $\mathbf{P}_1$ and $\mathbf{P}_2$ vectors of equation 10 are present. Although for lossless conductors, they are everywhere equal and opposite, it will be illuminating to concentrate for the moment on the $\mathbf{P}_2$ component alone, which, represents a continuation of the P_2^θ component in the gap, following its curvature near the surface of the stator.

Consider a typical kth conductor of Fig. 4, assumed rectangular in cross-section. The field $\mathbf{E}_2$ within it is still given by equation 6. The magnetic field of the conductor current, which from equations 11 to 13, and 19 may be written as

$$i(\tau,m)=I\sqrt{2}\cos(\tau-m\pi/3) \qquad (25)$$

is in the $-\theta$ direction and can be easily calculated to be

$$H_a^\theta(q,\tau,m)=-\frac{Z}{\pi Dh}qi(\tau,m) \qquad (26)$$

The above assumes closely spaced conductors and high permeability iron. The flow vector $\mathbf{P}_2$ is therefore directed radially upward, and is given by equations 8, 6, and 26, with the aid of equations 25 and 21 as

$$P_{2a}^r=-E_2^zH_a^\theta=\left(\frac{3VI}{\pi D\lambda}\right)\frac{q}{h}\times [S(\theta,m)+S'(\theta,\tau,m)] \qquad (27)$$

$$S(\theta,m)=\frac{\pi}{3}\cos(\theta-m\pi/3) \qquad (28)$$

$$S'(\theta,\tau,m)=\frac{\pi}{3}\cos\times[(\theta-m\pi/3)-2(\tau-m\pi/3)] \qquad (29)$$

The S functions are normalized forms of P_{2a}^r, with the factor $3VI/\pi D\lambda$ representing an average flow of the total machine power

$$p_0=3VI \qquad (30)$$

across the stator surface $\pi D\lambda$. $S(\theta,m)$ is independent of time, and represents a time average of P_{2a}^r. Once again periodicity with respect to the six belts exists, and an analysis per belt is sufficient, setting $m=0$ in equations 28 and 29. The $S(\theta)$ function is plotted on Fig. 3. It is apparent that the net transfer of rotor energy to the stator conductors is not uniform but is symmetric with respect to the center of the belt. Since the average value of $S(\theta)$ is unity, i.e.

$$S_{av}=\frac{3}{\pi}\int_{-\pi/6}^{\pi/6}S(\theta)d\theta=1 \qquad (31)$$

the total average power input to each belt is given by

$$\frac{3VI}{\pi D\lambda}S_{av}\frac{\pi D\lambda}{6}=\frac{p_0}{6} \qquad (32)$$

as a check on the procedure.

We note that P_{2a}^r, varies linearly with q within the conductors—a consequence of assumed uniform current distribution. Everywhere within the conductors, this energy is being continually converted to the equal and opposite P_{1a}^r energy; the rate of conversion, proportional to $\partial P_{2a}^r/\partial q$, being constant, with no net loss of energy. Had the conductor losses been included in the analysis, then the difference between the P_{2a}^r and P_{1a}^r flow would account for these losses, adding further meaning to the separation of energy flow into these two components. This extension of the analysis may easily be performed as was done in the d-c machine,[1] but will not be presented here. Inclusion of skin effect or more complex conductor geometry would require a corresponding modification of the $\mathbf{H}_a$ field within the conductors, resulting in more complexity in the expressions for energy flow vectors, without fundamentally altering the picture of energy flow.

The exact analysis of the transition region just below the surface of the stator, where the curvature of the $\mathbf{H}_a$ field bends the lines of P_2^θ flow in the gap until they become the lines of the P_{2a}^r flow at the conductors, cannot be easily done analytically, as it requires the complete knowledge of the magnetic field distribution, but it may be performed graphically.

Energy Flow at the Surface of the Rotor

As already noted, the energy flow vectors in the gap and at the stator conductors are neither uniform with the angular co-ordinate θ, nor is their variation with θ the same. This difference must, of course, be accounted for, in order that an energy balance be everywhere preserved, by the flow of energy at the rotor surface. While the exact specification of the fields there is quite complex, the resultant radial component of the flow vector P_{2b}^r may be at once found from the energy-balance requirement.

Consider a slice of the gap of thickness $Dd\theta/2$, height g, and length λ, shown on Fig. 5. In steady-state, instantaneous energy balance requires that

$$\lambda D d\theta/2[P_{2b}^r(\theta,\tau)-P_{2a}^r(\theta,\tau)]+\lambda g[P_2^\theta(\theta,\tau)-P_2^\theta(\theta+d\theta,\tau)]=0 \quad (33)$$

Consequently

$$P_{2b}^r=P_{2a}^r+(2g/D)(\partial P_2^\theta/\partial\theta) \quad (34)$$

Evaluating, with the aid of equations 22 to 24 and 27 to 29

$$P_{2b}^r=\frac{3VI}{\pi D\lambda}[R(\theta,m)+R'(\theta,\tau,m)] \quad (35)$$

$$R=\frac{\pi^2}{6\sqrt{3}}\cos(\theta-m\pi/3)+\frac{\pi}{3}(\theta-m\pi/3)\sin(\theta-m\pi/3) \quad (36)$$

$$R'=\frac{\pi^2}{6\sqrt{3}}\cos[(\theta-m\pi/3)-2(\tau-m\pi/3)]+\frac{\pi}{3}(\theta-m\pi/3)\sin[(\theta-m\pi/3)-2(\tau-m\pi/3)] \quad (37)$$

The R functions are, in analogy with the S functions of equations 28 and 29, normalized flow vector components, which are again periodic with respect to the six belts. Here again, R of equation 36 represents the time average of the P_{2b}^r vector component, and is plotted on Fig. 3 together with G and S. It is symmetric with respect to the center of the belt and its average over the belt is again unity.

An over-all picture of the average flow of energy from the rotor towards the stator can thus be obtained at a glance from Fig. 3. Energy flows radially out of the rotor, mostly at the boundaries between the phase belts when viewed from the stator. The flow is directed towards the center of the belt in the gap, curves and then flows radially into the rotor conductors, mostly in the central region of each belt. This picture of energy flow is consistent with the qualitative discussion of energy flow in a salient-pole alternator discussed by Slepian,[4] where it was demonstrated that the flow out of the rotor occurs mainly at the pole tips, corresponding to the boundaries of the assumed 2-phase windings.

It is of interest to remark further on the time-varying components, G', S', and R'. Letting for generality

$$\theta'=\theta-m\pi/3;\ \tau'=\tau-m\pi/3 \quad (38)$$

equations 37, 24, and 29 may be rewritten as

$$R'=-\frac{\pi^2}{6\sqrt{3}}\cos[\theta'+2(\tau'-\theta')]-\frac{\pi}{3}\theta'\sin[\theta'+2(\tau'-\theta')] \quad (39)$$

$$G'=\frac{\pi}{12\sqrt{3}}\sin[\theta'+2(\tau'-\theta')]-\frac{\theta'}{6}\cos[\theta'+2(\tau'-\theta')] \quad (40)$$

$$S'=\frac{\pi}{3}\cos[\theta'+2(\tau'-\theta')] \quad (41)$$

With respect to the rotor frame of reference (that is, letting $\theta=\tau$ in the foregoing) these expressions reduce to $-R$, $-G$, and $-S$ functions of equations 36, 23, and 28, respectively, indicating the invariance of energy flow with respect to the change of the frame of reference, as required.

Flow of Energy to the Output Terminals

The flow vector $\mathbf{P}_1$ of equation 10 will now be shown to redirect the energy, which crosses the gap into the stator conductors, out of the stator conductors and towards the machine output terminals. This flow is analogous to the flow of energy in d-c machines[1] out of the armature towards the output terminals, except for the time variation here.

Within the stator conductor, the flow vector P_1^r is, as mentioned previously, equal and opposite to P_{2a}^r of equation 27, except for lossy conductors, where the difference between them accounts for the internal ohmic loss. As the lines of the $\mathbf{P}_1$ vector emerge from the stator conductors, they curve as a result of the curvature of the $\mathbf{H}_a$ field, and because the $\mathbf{E}_1$ field of the stator charges acquires a θ-directed component. The $\mathbf{P}_1$ vector thus acquires θ-directed and z-directed components. Once again, because of the complexity of the field in the immediate vicinity of the conductors, we proceed to analyze the flow in the region of the gap sufficiently removed from the stator, past this transition region, where only these last two components are present, leaving the transition region to a graphical solution, or to an approximate graphical fill-in procedure.

In this region the $\mathbf{H}_a$ field is again given by equation 20, and the E_1^z field by the negative of equation 6 for lossless cen-

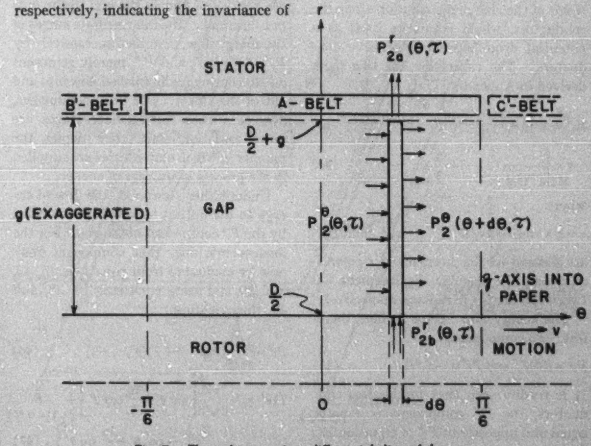


Fig. 5. Flow of energy in a differential slice of the gap

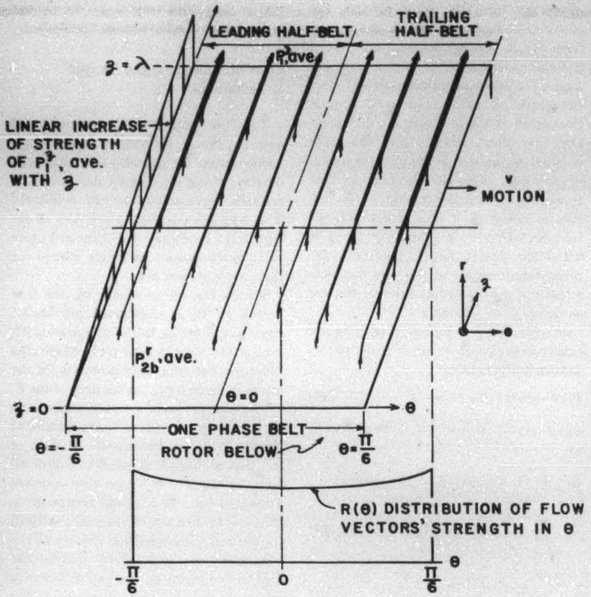


Fig. 6. Net flow of active energy in the gap. Pictorial representation

ductors. Using equation 21, this may be written

$$E_1^z(\theta, \tau) = -\frac{\sqrt{2}\pi V}{Z\lambda}\cos(\theta-\tau) \tag{42}$$

In addition, an $E_1^\theta(\theta, z, \tau)$ field appears, a result of the charge distribution along the conductors, which manifests itself as a potential drop between successive conductors. The expression for this field, derived in Appendix III, is

$$E_1^\theta(\theta, z, \tau) = \frac{2\sqrt{2}V}{D}\times \left[-\cos(\theta-\tau)+\frac{\delta u}{2}\sin(\theta-\tau)\right] \tag{43}$$

where

$$u = z/\lambda \text{ and } \delta = 2\pi/Z \tag{44}$$

are defined as the normalized z-co-ordinate and the angular displacement between conductors in radians, respectively.

The energy flow vector $\mathbf{P}_1$ has, therefore, two components

$$P_1^\theta = E_1^z H_a^r \text{ and } P_1^z = -E_1^\theta H_a^r \tag{45}$$

It is readily seen that for lossless conductors, the P_1^θ component is exactly equal and opposite to P_2^θ of equation 22. It appears, therefore, that, except for stator losses, it is superfluous to consider any of the studied components, except the P_{2b}^r flow of equation 35 and the P_1^z flow below. The upward flow of energy from the rotor, after appropriate curvature in the transition region, virtually becomes the z-directed flow in the gap to the output terminals; the intermediate stages of the energy flow, described successively by P_2^θ, P_{2a}^r, P_{1a}^r, and P_1^θ, merely represent an alternating undiminished flow into and out of the stator. For lossy conductors, however, these intermediate stages are quite essential, since they depict the manner in which stator losses are supplied in the process of transfer of energy.

Under either viewpoint, the flow of energy to the output terminals is described by the P_1^z component in the gap. For the lossless machine, this component may now be evaluated from equations 45, 43, and 20, and using equations 18, 38, and 44, thus yielding

$$P_1^z = \frac{3VI}{\pi Dg}[T(\theta', z) + T'(\theta', z, \tau')] \tag{46}$$

$$T(\theta', z) = \frac{\pi Z}{6\sqrt{3}}\left[\sin\theta' + \frac{\delta u}{2}\cos\theta'\right] + \frac{Z}{3}\theta'\left[-\cos\theta' + \frac{\delta u}{2}\sin\theta'\right] \tag{47}$$

$$T'(\theta', z, \tau') = \frac{\pi Z}{6\sqrt{3}}\left[-\sin(\theta'-2\tau') - \frac{\delta u}{2}\cos(\theta'-2\tau')\right] + \frac{Z}{3}\theta'\times \left[-\cos(\theta'-2\tau') + \frac{\delta u}{2}\sin(\theta'-2\tau')\right] \tag{48}$$

Periodicity of the above expressions with respect to the phase belts is again apparent, and it will suffice to consider the A-belt only, where $\theta'=\theta$. The average power output from a differential slab of the gap of area $g\lambda$ and width $Dd\theta/2$ is

$$dp = (gDd\theta/2)\left(\frac{3VI}{\pi Dg}\right)[T(\theta,\lambda) - T(\theta, 0)] \tag{49}$$

$$= (p_0/6)\left[\frac{\pi}{2\sqrt{3}}\cos\theta + \theta\sin\theta\right]d\theta \tag{50}$$

Since

$$\int_{-\pi/6}^{\pi/6}\left[\frac{\pi}{2\sqrt{3}}\cos\theta + \theta\sin\theta\right]d\theta = 1 \tag{51}$$

a check on an over-all energy balance is obtained. Furthermore, comparison with equations 35 and 36 shows that this differential power flow dp is

$$dp = \frac{3VI}{\pi D\lambda}R(\theta)(D\lambda d\theta/2) \tag{52}$$

indicating that an energy balance holds for each differential slab of the gap, with the P_{2b}^r flow eventually becoming the P_1^z flow in the gap. This check is particularly gratifying, since these flow vectors were obtained by different methods.

Consider the T function of equation 47 in more detail. It may be conveniently looked upon as composed of two parts, one independent of z, and one varying linearly with z

$$T(\theta, z) = T_1(\theta) + uT_2(\theta) \tag{53}$$

$$T_1(\theta) = \frac{Z}{3}\left[\frac{\pi}{2\sqrt{3}}\sin\theta - \theta\cos\theta\right] = 2ZG(\theta) \tag{54}$$

$$T_2(\theta) = \frac{\pi}{3}\left[\frac{\pi}{2\sqrt{3}}\cos\theta + \theta\sin\theta\right] = R(\theta) \tag{55}$$

The T_1 function corresponds closely to the G function of equation 23, the average component of the P_2^θ flow vector. Its average over the belt is zero, and it represents energy that is circulating within the gap, in the direction of the rotor axis in the trailing half space of the belt, and in an opposite direction in the leading half space. (These directions of circulatory flow in the gap are reversed if the sense of phase belt windings with respect to the sense of rotation is opposite to that assumed here.) The amount of this circulating energy is directly proportional to the number of conductors Z. The T_2 function on the other hand, independent of Z, represents the net flow of energy out of the machine. Its functional dependence on θ is the same as that of the R function of equation 36, to which it is identically equal, as shown plotted on Fig. 3. Thus, the axial average net-energy flow vector in the gap has the same distribution in θ as the radial average energy flow vector at the rotor surface, starting from zero at one end of the machine and increasing linearly with z over the axial length λ to its maximum at the other end surface, by virtue of the coming energy from the rotor below. The situation is depicted on Fig. 6.

It remains to indicate how this P_1^z-directed flow, or, specifically the T_2 flow, reaches the machine output terminals. Fig. 7 shows qualitatively the conditions that prevail when one complete phase (belts A and A') is viewed as a whole. The Maxwellian potential drop between the A and A' conductors sets up a leakage field $\mathbf{E}_1'$, which, together with the magnetic leakage field $\mathbf{H}'$ results in a transmission-line-like flow of energy along this A—A' "transmission line," as discussed by Slepian,[5] and described by the $\mathbf{P}'=\mathbf{E}_1'\times\mathbf{H}'$ flow vector. Analytical treatment of this flow, including, for example, the space curvature of the $\mathbf{P}_1^z$ lines into the $\mathbf{P}'$ lines along this "transmission line," is beyond the scope of this paper, but it should be clear from the fundamental analysis underlying the concept of the V-i postulate in the reference cited that an over-all energy balance is at all times preserved in the process.

Flow of Energy Under Arbitrary Load Condition

Generalization of the previous analysis to any load may now be made. The essence of this generalization lies in the fact that, in general, the current in the stator conductors may have an algebraic phase lag of β radians (lead corresponding to negative values of β) with respect to the induced electromotive force proportional to E_2^z or to H_b^r. The simplest approach is to retain the expressions for the currents and their fields of equations 25, 20, and 26, and to modify the corresponding expressions for the fields $\mathbf{H}_b$, $\mathbf{E}$ and the power output to read

$$p_0=3VI\cos\beta \tag{56}$$

$$-E_1^z=E_2^z=\frac{2\pi V}{Z\lambda}\cos(\theta-\tau-\beta) \tag{57}$$

$$E_1^\theta=\frac{2\sqrt{2}V}{D}\left[-\cos(\theta-\tau-\beta)+\frac{\delta u}{2}\times\right.$$
$$\left.\sin(\theta-\tau-\beta)\right] \tag{58}$$

The corresponding physical interpretation is that now at the instant when the current in a belt is maximum, the magnetomotive-force wave of the rotor will have moved forward by an angular displacement of β radians with respect to the center of the belt.

It is convenient to expand the components of **E** into in-phase (active) and out-of-phase (reactive) components, designated by subscripts c and s

$$-E_1^z=E_2^z=E_{2c}^z+E_{2s}^z \tag{59}$$

$$E_1^\theta=E_{1c}^\theta+E_{1s}^\theta \tag{60}$$

$$E_{2c}^z=\frac{\delta V}{\sqrt{2}\lambda}\cos\beta\cos(\theta-\tau) \tag{61}$$

$$E_{2s}^z=\frac{\delta V}{\sqrt{2}\lambda}\sin\beta\sin(\theta-\tau) \tag{62}$$

$$E_{1c}^\theta=\frac{2\sqrt{2}V}{D}\cos\beta\times$$
$$\left[-\cos(\theta-\tau)+\frac{\delta u}{2}\sin(\theta-\tau)\right] \tag{63}$$

$$E_{1s}^\theta=\frac{2\sqrt{2}V}{D}\sin\beta\times$$
$$\left[-\sin(\theta-\tau)-\frac{\delta u}{2}\cos(\theta-\tau)\right] \tag{64}$$

The active c components combine with the $\mathbf{H}_a$ field yielding the active-flow vector components which are identical with those previously described under unity-power-factor load, except for a multiplying factor of $\cos\beta$ throughout, which, in view of equation 56, is consistent with the requirement of energy balance. The reactive s components are therefore expected to result in zero net energy transfer. Their derivation, which is analogous to the previous procedure, yields

$$P_{2s}^\theta=-\left(\frac{3VI}{\lambda g}\right)\frac{\sin\beta}{2\pi}\times$$
$$[R(\theta,m)+R'(\theta,\tau,m)] \tag{65}$$

$$P_{2as}^r=\left(\frac{3VI}{\pi D\lambda}\right)\sin\beta\times$$
$$\left[\frac{\pi}{3}\sin\theta'+\frac{\pi}{3}\sin(\theta'-2\tau')\right] \tag{66}$$

$$P_{2bs}^r=\left(\frac{3VI}{\pi D\lambda}\right)2\pi\sin\beta\times$$
$$[G(\theta,m)+G'(\theta,\tau,m)] \tag{67}$$

$$P_{1s}^r=-P_{2as}^r(\text{lossless stator}) \tag{68}$$

$$P_{1s}^\theta=-P_{2s}^\theta(\text{lossless stator}) \tag{69}$$

$$P_{1s}^z=\frac{3VI}{\pi Dg}\sin\beta\times$$
$$[U(\theta,z,m)+U'(\theta,z,\tau,m)] \tag{70}$$

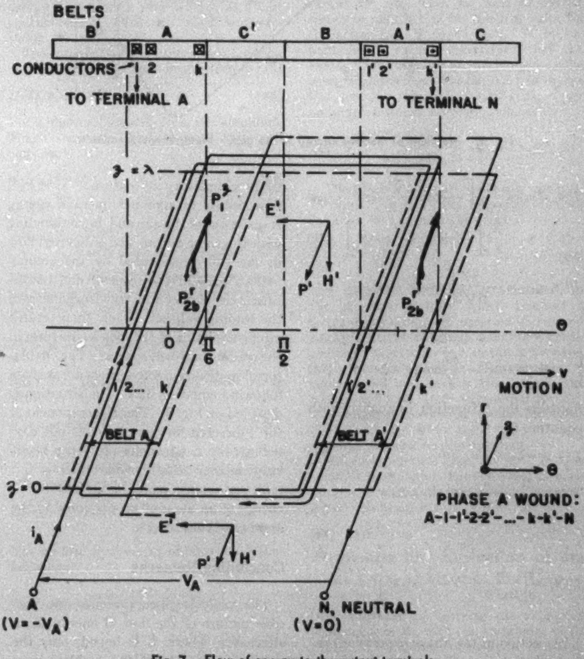


Fig. 7. Flow of energy to the output terminals

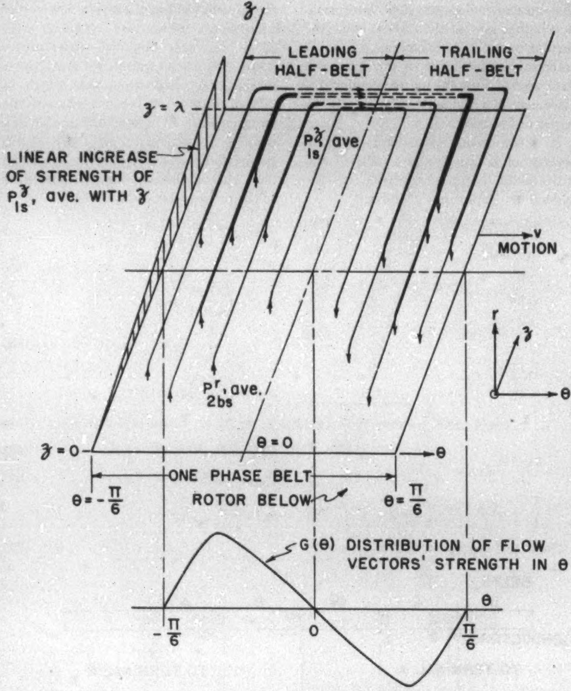


Fig. 8. Net flow of reactive energy in the gap. Pictorial representation

$$U(\theta, z, m)=\frac{\pi Z}{6\sqrt{3}}\left[-\cos\theta'+\frac{\delta u}{2}\sin\theta'\right]+\frac{Z}{3}\theta'\left[-\sin\theta'-\frac{\delta u}{2}\cos\theta'\right] \quad (71)$$

$$U'(\theta, z, \tau, m)=\frac{\pi Z}{6\sqrt{3}}\left[\cos(\theta'-2\tau')-\frac{\delta u}{2}\sin(\theta'-2\tau')\right]+\frac{Z}{3}\theta'\left[-\sin(\theta'-2\tau')-\frac{\delta u}{2}\cos(\theta'-2\tau')\right] \quad (72)$$

Splitting the U function, in analogy with equation 53

$$U(\theta, z)=U_1(\theta)+u\,U_2(\theta) \quad (73)$$

$$U_1(\theta)=\frac{Z}{3}\left[-\frac{\pi}{2\sqrt{3}}\cos\theta'-\theta\sin\theta'\right]=-\frac{Z}{\pi}R(\theta) \quad (74)$$

$$U_2(\theta)=\frac{\pi}{3}\left[\frac{\pi}{2\sqrt{3}}\sin\theta'-\theta'\cos\theta'\right]=2\pi G(\theta) \quad (75)$$

Inspection of the above expressions reveals that the reactive energy flow is entirely circulatory in nature. The $P_{2s}{}^{\theta}$, flow results in zero net average energy transfer out of the gap and, in conjunction with the rotor output and z-directed flow in the gap, represented by the components $P_{2bs}{}^{r}$ and U_2, respectively, constitutes a circulatory flow, which originates in the leading half-belt space, then returns after completion of the loop to the rotor in the trailing half-belt space. (The circulatory flow direction conventions are as given following equation 55.) The situation is depicted on Fig. 8. The U_1 component of the z-directed flow vector does not contribute any net transfer of energy by its very nature, being constant with z, but represents once again a circulation of energy of an amount proportional to the number of conductors.

Concluding Remarks

The analysis given presents one possible picture of the flow of energy in an alternator, where, it is hoped, only the significant aspects of the problem were outlined. The picture is far from a complete one, since only a somewhat idealized alternator has been treated. While an extension of the analysis in certain cases is fairly straightforward, an attempt to remove several assumptions would constitute a formidable task. Inclusion of conductor losses has already been discussed here, and outlined in a previous paper on d-c machines[1]; it presents no special difficulty. Variation with radial distance for an appreciably thick gap is considered in Appendix III. Extension of the analysis to synchronous motors is likewise straightforward.

The refinement of the foregoing analysis to take into account such practical effects as saliency, nonuniform gap, the presence of slots and ventilating ducts, a finite number of conductors of practical size and shape, or relatively low-permeability iron with, perhaps, appreciable losses, entails serious computational difficulties and requires the exact specification of the fields, which must be obtained by a point-to-point field plot or by measurement. Such information is, indeed, required even in the idealized machine for the electric field in the transition regions, discussed in the paper, for the complete specification of the flow vectors there. Elimination of the assumption of linearity to include saturation and hysteresis effects presents a most serious conceptual as well as computational obstacle. This has not been attempted. At worst, the analysis based on these latter assumptions may be only deemed valid for lightly loaded machines at low voltage (low speed or low excitation).

In spite of these shortcomings, it is hoped that the study of energy flow presented will serve a useful purpose of adding to the basic understanding of the operation of synchronous machines, and, as such, may be of some potential value to the designer of practical machines. It is furthermore hoped that this analysis will serve as the basis from which similar studies of energy flow in asynchronous machines or commutator-type a-c machines may evolve.

Appendix I. An Approximation Procedure for Poynting's Vector at Low Frequencies

If an electromagnetic field is produced by two (extension to any number is apparent) independent sources (stationary and/or moving charges), a and b, then the resultant Poynting's vector has the following components

$$\mathbf{P}=\mathbf{P}_a+\mathbf{P}_b+\mathbf{P}_{ab}+\mathbf{P}_{ba}=(\mathbf{E}_a+\mathbf{E}_b)\times(\mathbf{H}_a+\mathbf{H}_b) \quad (76)$$

$$\mathbf{P}_a=\mathbf{E}_a\times\mathbf{H}_a;\quad \mathbf{P}_b=\mathbf{E}_b\times\mathbf{H}_b \tag{77}$$

$$\mathbf{P}_{ab}=\mathbf{E}_a\times\mathbf{H}_b;\quad \mathbf{P}_{ba}=\mathbf{E}_b\times\mathbf{H}_a \tag{78}$$

It is worth while to investigate the conditions under which the cross components $\mathbf{P}_{ab}$ and $\mathbf{P}_{ba}$ contribute negligibly to the overall flow of energy. This situation results in a considerable simplification of the picture of energy flow in rotating devices. Its criterion is a negligible small value of divergence of the sum $\mathbf{P}_{ab}+\mathbf{P}_{ba}$ in comparison with that of $\mathbf{P}_a$ or $\mathbf{P}_b$ or both. Expanding

$$\operatorname{div}(\mathbf{P}_{ab}+\mathbf{P}_{ba})=\mathbf{H}_a\cdot\operatorname{curl}\mathbf{E}_b-\mathbf{E}_b\cdot\operatorname{curl}\mathbf{H}_a+\mathbf{H}_b\cdot\operatorname{curl}\mathbf{E}_a-\mathbf{E}_a\cdot\operatorname{curl}\mathbf{H}_b \tag{79}$$

Equation 79 is identically zero in the most general case (in addition to several trivial cases) when curl $\mathbf{E}_a=0$ $\mathbf{H}_a=0$ (that is, $\mathbf{E}_a$ is an electrostatic field) and $\mathbf{H}_b$ is a magnetostatic field produced by steady currents b. (See appendix of reference 1. Some care must be exercised in applying this. If the b conductors exist in polarizable media, for instance, then within them $\mathbf{E}_a$ is no longer zero and div $(\mathbf{P}_{ab}+\mathbf{P}_{ba})\neq 0$. Subscripts a and b are, of course, reversible in this discussion.)

A limited number of practical cases exists where the above requirements are satisfied. It was shown in a previous paper[1] how the study of energy flow in one such case—the d-c machine—is simplified by ignoring these cross-components of **P**. Commonly quoted "paradoxes" of energy flow where no energy apparently should be flowing always involve a circulatory flow of energy, which can be similarly traced to these divergenceless cross components of **P**.[3] It is now of interest to investigate a wider class of occurrences, where certain of these cross components are not zero, but are still negligibly small. The a-c synchronous machine falls into this category. It is therefore desirable to establish a criterion for the validity of neglecting such components in order to achieve a possible simplification in the study of the flow of energy.

Consider the situation where curl $\mathbf{E}_a$ is still zero, but $\mathbf{H}_b$ varies slowly with time, associated with low-frequency currents in the b conductors. Equation 79 then becomes

$$\operatorname{div}(\mathbf{P}_{ab}+\mathbf{P}_{ba})=-\mathbf{E}_a\cdot\operatorname{curl}\mathbf{H}_b \tag{80}$$

Using one of Maxwell's field equations[1]

$$\operatorname{curl}\mathbf{H}_b=\mathbf{J}_b+\epsilon\partial\mathbf{E}_b/\partial t \tag{81}$$

equation 80 may be restated to read

$$\operatorname{div}(\mathbf{P}_{ab}+\mathbf{P}_{ba})=0 \tag{82}$$

within the b conductors where **E** is zero, and

$$\operatorname{div}(\mathbf{P}_{ab}+\mathbf{P}_{ba})=\epsilon_0\mathbf{E}_a\cdot\partial\mathbf{E}_b/\partial t \tag{83}$$

outside of the b conductors. This last equation represents the error involved in ignoring the cross components of **P**, corresponding to the energy required to maintain the $\mathbf{E}_a$ field in the presence of the displacement currents due to the time variation of the $\mathbf{H}_b$ field. This energy should be compared with the energy supplied by the b sources for the maintenance of the b field, whose flow is described by $\mathbf{P}_b$. It will suffice to compare a component of div $\mathbf{P}_b$ with equation 83, namely

$$-\mathbf{E}_b\cdot\operatorname{curl}\mathbf{H}_b=\mathbf{E}_b\cdot(\mathbf{J}_b+\epsilon_0\partial\mathbf{E}_b/\partial t) \tag{84}$$

If the $\mathbf{E}_a$ and $\mathbf{E}_b$ fields are in the same direction and of the same order of magnitude, or even if they differ by a factor as large as 10^3, then for frequencies of the order of power frequencies the term $\mathbf{E}_b\cdot\mathbf{J}_b$ is much larger than either $\mathbf{E}_b\cdot\epsilon_0\partial\mathbf{E}_b/\partial t$ or $\mathbf{E}_a\cdot\epsilon_0\partial\mathbf{E}_b/\partial t$, suggesting that the cross component $\mathbf{P}_{ab}$ may be ignored in the study of energy flow.

In the a-c machine of this paper both the a fields of the stator conductors and the b fields of the rotor are time varying. It is still possible, however, to use the conclusions of the foregoing discussion to advantage by separating the $\mathbf{E}_a$ field into two components, $\mathbf{E}_1$ and $\mathbf{E}_2$, in such a useful manner, as defined by equations 4 and 5, that the $\mathbf{E}_1$ field (set up by the stator charges) has no curl. The cross component $\mathbf{E}_1\times\mathbf{H}_b$ can then be neglected in comparison with other components, such as $\mathbf{E}_1\times\mathbf{H}_a$, or $\mathbf{E}_2\times\mathbf{H}_b$, as evident from the previous discussion. This approximation is even better than the previous comparison with the $\mathbf{E}_b\times\mathbf{H}_b$ component indicated.

Appendix II. Variation with Radial Distance in the Gap

While in the main body of the paper, it has been assumed that the gap width is negligibly small, it is of interest to investigate the manner in which the fields and the energy flow vectors vary with r, as a first approximation, when this width is appreciable.

The rotor field will in general be a function of r, and equation 1 must be modified to read

$$H_b'(r,\theta,\tau)=H'(r)\cos(\theta-\tau) \tag{85}$$

For a high-permeability iron, the amplitude function may be obtained from Maxwell's equation[1]

$$\operatorname{div}(\mu_0\mathbf{H})=0 \tag{86}$$

as follows

$$\mu_0H_b^r(r,\theta,\tau)r\,d\theta dz=\mu_0H_b^r(D/2,\theta,\tau)(D/2)d\theta dz \tag{87}$$

Using the symbol H' of equation 1 to designate the amplitude of H_b^r at the rotor surface, $r=D/2$, we obtain

$$H_b^r(r,\theta,\tau)=\frac{DH'}{2r}\cos(\theta-\tau) \tag{88}$$

Since the velocity of the H_b^r field varies with r

$$v(r)=\omega r=2vr/D \tag{89}$$

where v is the rotor surface velocity, the E_1^z and E_2^z fields are independent of r, and equation 6 is valid

$$E_2^z=-E_1^z=\mu_0v(r)H_b^r(r,\theta,\tau)=\mu_0vH'\cos(\theta-\tau) \tag{90}$$

The no-load flow vector of equation 9 thus becomes

$$P_{2b}^\theta(r,\theta,\tau)=\frac{\mu_0\omega D(H')^2}{4}\left(\frac{D}{2r}\right)^2\times[1+\cos(2\theta-2\tau)] \tag{91}$$

and its divergence

$$(1/r)(\partial P_{2b}^\theta/\partial\theta)=-\mu_0\omega(H')^2\left(\frac{D}{2r}\right)^2\sin(2\theta-2\tau) \tag{92}$$

still has zero average value.

The modifications for a loaded machine will be restricted to a unity-power-factor load, the conclusions applying equally to the reactive components of the flow vectors. The modification required first is in the expression for H of equation 18, since functional dependence on θ of the $\mathbf{H}_a$ field is unaffected. The amplitude H is in general a function of r, and may be calculated from Ampere's law

$$\oint\mathbf{H}\cdot d\mathbf{s}=NI \tag{93}$$

for high permeability iron as follows

$$2\int_{D/2}^{D/2+g}H(r)dr=ZI(\sqrt{2}/6) \tag{94}$$

Invoking the requirement of equation 86 again, the above may be solved for $H(r)$ to yield

$$H(r)=\frac{D}{2r}\,\frac{\alpha}{\log(1+\alpha)}\,\frac{ZI\sqrt{2}}{12g} \tag{95}$$

where

$$\alpha=2g/D \tag{96}$$

Equation 95 reduces to the previous expression of equation 18 for small values of α and $r=D/2$.

The modification required of the flow vector P_2^θ may then be seen to consist merely of multiplying the previous value by the factor $(D\alpha)/[2r\log(1+\alpha)]$. It can be, furthermore, readily ascertained that the values of P_{2a}^r and P_{2b}^r are unchanged. The energy-balance equation 33, used to derive P_{2b}^r, now has the product $\lambda_gP_2^\theta$ replaced by an integral, which, however, reduces to

$$\lambda P_2^\theta\int_{D/2}^{(D/2)+g}\frac{D}{2r}\,\frac{\alpha}{\log(1+\alpha)}dr=\lambda_gP_2^\theta \tag{97}$$

The effect of an appreciable gap width on E_1^θ and P_1^z arises from reconsideration of equation 99 in which the term

$$\frac{\partial E_1^\theta}{\partial r}+\frac{E_1^\theta}{r}=0 \tag{98}$$

now makes an additional appearance. In the previous analysis, which assumed a (rectangular) thin gap, equation 98 reduced to the identity: E_1^θ independent of r, corresponding to setting $r=\infty$ in the E_1^θ/r term, implicit in the analysis. Examination of equation 98 now reveals that, again, the correction factor $D/2r$ is required for E_1^θ. The required correction factors of order $(D/2r)$ and $(D/2r)^2$ for the P_1^θ and P_1^z components, are now readily apparent.

Since for practical gaps, the ratio $2g/D$ is small, it is apparent that the effect of variation with radial distance r in the gap is of minor significance.

Appendix III. Evaluation of the $E_1^\theta(\theta, z, \tau)$ Field in the Gap

The method of deriving the θ-directed component of $\mathbf{E}_1$, with the knowledge of its z-directed component of equation 42, parallels the method employed in the analysis of the d-c machine.[1] It is based on the fact that

$$\text{curl } \mathbf{E}_1 = 0 \tag{99}$$

which here yields

$$\frac{\partial}{\partial z} E_1^{\theta} = \frac{2}{D} \frac{\partial}{\partial \theta} E_{1z} \tag{100}$$

and on the knowledge of the variation of E_1^{θ} with θ at the boundary $z=0$. This variation is such as to account for the voltage drop between successive conductors, assumed here to be closely spaced, which is also equal to twice the drop along the conductors at corresponding θ, associated with the E_1^s field within them. The factor of two accounts for the conductor 180 degrees away. Thus

$$\left(\frac{\pi D}{Z}\right) E_1^{\theta}(\theta, 0, \tau) = 2\lambda E_1^{s}(\theta, \tau) \tag{101}$$

Noting that E_1^s, and hence $\partial E_1^s/\partial\theta$, are both independent of z, we may write on the basis of equation 100

$$E_1^{\theta}(\theta, z, \tau) = E_1^{\theta}(\theta, 0, \tau) + \int_0^{z} \left(\frac{2}{D}\right)(\partial E_1^{s}/\partial\theta)dz \tag{102}$$

$$= E_1^{\theta}(\theta, 0, \tau) + \left(\frac{2z}{D}\right)(\partial E_1^{s}/\partial\theta) \tag{103}$$

Evaluating the foregoing, and using equations 42 and 101, and the definitions of equation 44, equation 43 is readily obtained.

References

1. Flow of Energy in Direct Current Machines, **Edward I. Hawthorne.** *AIEE Transactions*, vol. 72, pt. I, Sept. 1953, pp. 438–45.

2. *Ibid.* For detailed discussion, see section on "Energy and Energy Flow Postulates," pp. 438–40, and references cited there.

3. Electric Machinery. Vol. II—Alternating Current Machines (book), **M. Liwschitz-Garik, C. C. Whipple.** D. Van Nostrand Company, Inc., New York, N. Y., 1946, section 5-2, p. 129.

4. The Flow of Power in Electrical Machines, **J. Slepian.** *The Electric Journal*, East Pittsburgh, Pa., vol. XVI, no. 5, 1919, pp. 301-11.

5. Energy Flow in Electrical Systems—the VI Energy-Flow Postulate, **J. Slepian.** *AIEE Transactions (Electrical Engineering)*, vol. 61, Dec. 1942, pp. 835–41.

No Discussion

A Step Forward in Printing Telegraphy

A. S. BENJAMIN
MEMBER AIEE

W. J. ZENNER
NONMEMBER AIEE

Synopsis: Recent commercial and military printing telegraph communication requirements could not be fully met in an efficient and economical manner with existing equipment. This paper describes the objectives being reached in the development of a new integrated line of Teletype apparatus especially designed to adequately serve present and future printing telegraph applications and particularly covers the Teletype Model 28 Direct Keyboard Page Printer Set, Fig. 1. This set consists of a page typing unit, keyboard, electrical service unit, and their motor drive and cabinet housing. The other units which will soon be available are the tape perforator, tape typing unit, and transmitter distributor, which with the keyboard and page typing unit are so designed as to be readily arranged in many combinations to meet specific applications. Fig. 2 indicates the basic units and a few representative combinations.

History of Printing Telegraphy

COMMERCIAL telegrams were transmitted using printing telegraph machines as early as 1851, seven years after the commercial introduction of Morse. Modern printing telegraphy got its start in 1910, with the invention by Howard L. Krum[1] of the start-stop method of synchronization applied to printing telegraphy. Commercial success came in the early twenties, at which time such equipment was primarily used by the commercial telegraph companies and press associations. In 1930 printing telegraph moved rapidly into the business field, and since 1940 we have witnessed the extensive use of that form of communication in military service.

The original applications presented no great problems. Direct circuits were set up between telegraph offices or between press sending stations and receivers located in newspaper offices. Forty-word per minute speeds were considered adequate, messages were printed on plain paper fed from a roll or on individual blanks inserted by operators, and due to the fact that equipment was concentrated in largecenters, only a relatively small number of maintenance men were required.

As the business-field applications grew, the speed went up to 60 words per minute and new requirements were introduced, which resulted in the development of various accessories which were attached to existing units or installed separately. Since the original design of the units did not anticipate many of these functions, the parts for a given mechanism were placed wherever there was space with the result that the equipment was somewhat difficult to maintain. These added features included sprocket feed, horizontal and vertical tabs, motor control, contacts for various switching and control purposes, etc. The space for these mechanisms was limited and it soon became necessary to use external units to provide the multiplicity of control functions.

The 60-word-per-minute speed became too slow for certain applications and the equipment was modified for 75-word operation. Attempts were made to increase the speed to 100 words; however, high maintenance costs severely limited this use. The installation of printing telegraph equipment on ships, planes and

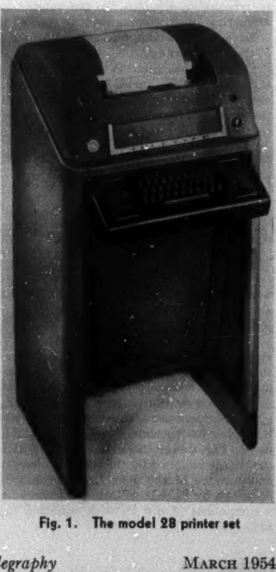


Fig. 1. The model 28 printer set

Paper 54-115, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 21, 1953; made available for printing December 4, 1953.

A. S. Benjamin and W. J. Zenner are with the Teletype Corporation, Chicago, Ill.

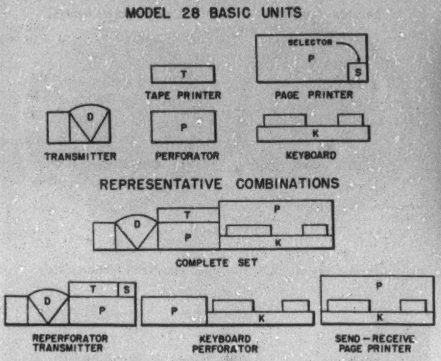

MODEL 28 BASIC UNITS
T
TAPE PRINTER
SELECTOR
P
S
PAGE PRINTER
D
TRANSMITTER
P
PERFORATOR
K
KEYBOARD
REPRESENTATIVE COMBINATIONS
D
T
P
P
K
COMPLETE SET
D
T
S
P
REPERFORATOR TRANSMITTER
P
K
KEYBOARD PERFORATOR
P
K
SEND—RECEIVE PAGE PRINTER

Fig. 2. Model 28 basic units and representative combinations

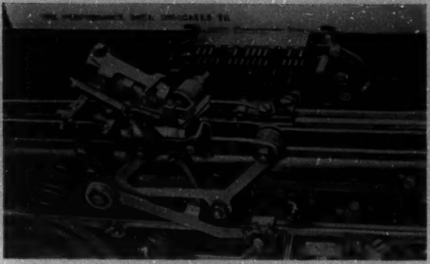


Fig. 4. When the typebox is in the desired position, the printing hammer drives the type pallet against the ribbon and paper

trucks imposed requirements of satisfactory operation at various inclinations and under conditions of great vibration.

As business expanded the use of printing telegraph techniques, the systems became larger and more complicated. Automatic switching was introduced and printers were installed in hundreds of offices of a single user. The Bell System Teletypewriter Exchange Service (TWX) which is similar to telephone exchange service, except that printers are connected through switchboards for typewritten communication, now has more than 35,000 Teletype sets distributed over the entire country. Each of these fields of use required new types of equipment. In almost all cases the equipment was wanted in a hurry, and since the future demand was uncertain, each new piece of equipment was generally provided by modifying an existing machine. This obviously did not result in an orderly development of standardized equipment.

Objectives of New Model 28 Line of Equipment

It became evident that a new line of equipment was needed which would meet the following prime objectives:

1. An integrated group of component units
2. Lower maintenance costs
3. Quieter operation
4. Increased speed.
5. Greater capacity for additional functions
6. Decreased weight

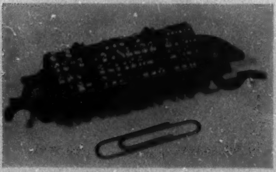


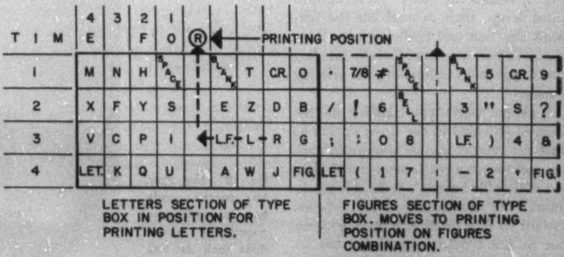
Fig. 3. The typebox, shown in comparison with an ordinary clip

A review of the typing mechanisms used in Teletype units of various kinds and those developed by other printing telegraph producers indicated that they were generally unsuited to meet the new requirements. Typebar carriages were considered too heavy and typewheels difficult to index at high speeds. Existing clutch designs similarly were considered inadequate; felt clutches having always been troublesome due to heat and loss of lubricant and positive clutches subject to breakage due to impact. It was therefore decided to design new elements to specifically meet the requirements of today and the next 10 to 20 years.

Some of the New Elements of the Model 28 Page Printer Set

The printing mechanism is the most obvious of the several improvments in the new machine. The type pallets, instead of being mounted on bars in the conventional manner are carried in a small rectangular box about 1/2 inch thick, 1 inch wide and 2 inches long, Fig. 3. Sixty-four pallets are arranged in four horizontal rows, each row having a capacity of 16 characters. To type a character, the typebox is moved to bring the desired character to the printing point, and a printing hammer, shown in Fig. 4, operates to drive the type pallets against the typewriter ribbon and paper. Each pallet is provided with a return spring that restores the pallet to its normal position after printing. After the printing hammer has operated, the typebox returns to its initial position below the printed line on the paper, so that the typing becomes visible.

Characters in the left half of the box are letters; those in the right half are figures. A shift mechanism is used to change from letters to figures. The normal position of the typebox for letters



T I M	4 E	3	2 F	1 O	R													
1	M	N	H	SPACE		BLANK	T	C.R.	O	·	7/8	#	SPACE		BLANK	5	C.R.	9
2	X	F	Y	S		E	Z	D	B	/	!	6	BELL		3	"	S	?
3	V	C	P	I	←	L.F. ←	L ←	R	G	;	:	0	8		L.F.	)	4	&
4	LET.	K	Q	U		A	W	J	FIG.	LET.	(	1	7		–	2	'	FIG.

PRINTING POSITION
LETTERS SECTION OF TYPE BOX IN POSITION FOR PRINTING LETTERS.
FIGURES SECTION OF TYPE BOX. MOVES TO PRINTING POSITION ON FIGURES COMBINATION.

Fig. 5. Typebox arangement and printing method

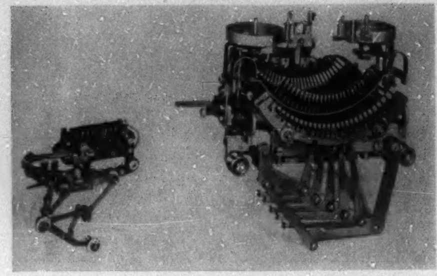


Fig. 6. The new typebox assembly weighs only a tenth of the older moving type basket carriage

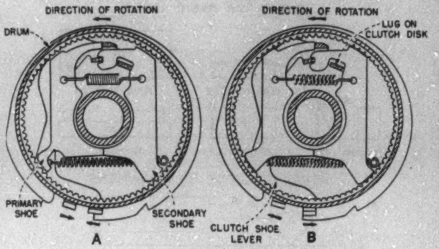


Fig. 7. The clutch is an internal expansion friction clutch

A. Disengaged
B. Engaged

selection is as shown in Fig. 5. When figures are to be printed, the centerline of the figures group is moved to the printing position. Movement of the typebox, in selecting the desired character to be printed, is controlled by two index mechanisms, one controlling the vertical motion to select the proper row of type, and the second controlling the horizontal motion to select the desired character in that row. These motions together form a rectangular co-ordinate system for all the 32 permutations of the standard five unit telegraph code.

The five elements of the code are used to position the pallet in a different manner than on other telegraph apparatus units. The first two elements of the code are used to determine the vertical position of the box; in other words, the nos. 1 and 2 impulses, by their presence or absence in the code permutation, will cause the printer to select one of four levels on the typebox for printing. The no. 3 impulse determines which way the box will move from either the letters or figures centerlines; that is, either to the right or left. The nos. 4 and 5 impulses determine how far the box will move in the selected direction; there are four locations in each. The code combination for *R*, as shown in Fig. 5, causes the box to move upward three levels, then to establish the leftward direction, and finally to move three spaces in that direction, thereby bringing the *R* type pallet in line with the printing hammer.

A unique toggle-type coupling mechanism is provided in the drive system so that the typebox can be stopped in various positions in a gentle manner and without noticeable impact. With this mechanism, the movement of the typebox toward its final position is at high speed, but as the typebox approaches the selected position, the toggle mechanism reduces its speed. At the end of the typebox travel, where further motion is blocked by the index mechanism, the speed of the typebox is about one fifth of that at which it would have passed this position.

Elimination of the conventional type basket greatly reduces the size and weight of the carriage that travels back and forth across the page, starting and stopping for each character printed. A comparison of the old and new type assemblies is illustrated in Fig. 6. In a standard model 15 typing unit, the moving carriage assembly weighs slightly over 5 pounds; the carriage assembly in the model 28 has a total weight of 8 ounces. This 10-to-1 weight reduction results in a very fast carriage return, the carriage returning easily within time of two character intervals (signals received for carriage return and paper feed) at 100 words per minute. Gravity has no significant effect on the carriage motion. The machine can be operated safely on shipboard or in other locations where it might not be in a level position.

Since the pallets are carried in an accurately made box instead of at the ends of the type bars, as in the older machine, type alignment is controlled by manufacture of the parts with no further adjustment required. Stability of alignment, too, is greater, and this should reduce maintenance. This small lightweight carriage also makes possible an over-all reduction in size and weight of the machine, and permits the use of stamped sheet-metal framing instead of the massive cast framework used on the older machines. The weight of the model 28 Teletype printer set is 38 pounds, not including the cabinet and accessories.

At the right end of the typebox, a small manually operated clamping lever holds the box in its supporting carriage. This clamping lever may be released with a light finger motion so that the typebox can be withdrawn from the machine without using tools. Thus the box may easily be cleaned without brushing dirt into the mechanism of the machine. Also, the typebox may be exchanged in a matter of seconds for another having different character faces. On older types of equipment with individual type, it is necessary to solder type pallets to typebars and realign them.

Since each character is separately

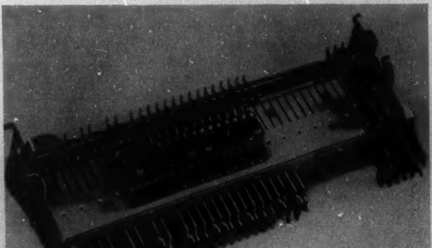


Fig. 8. The stunt box contains code bars for performing nonprinting operations such as line feed, carriage return, shift, etc.

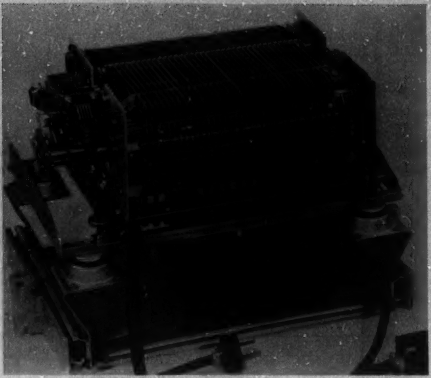


Fig. 9 (left). The sequential selector, which is replaced by the stunt box in the model 28 set

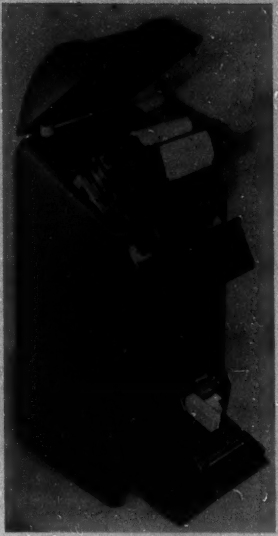


Fig. 10. The equipment is pivotally mounted in the cabinet and swings outward for maintenance. Incoming wires are connected to terminal blocks behind the typing unit

mounted on an individual type pallet, only one operates at a time. In the older machines, with two characters on a single pallet, there is a possibility of printing part of the other character on a pallet if the adjustment is not exact, or if one of the characters is worn, or if several carbon copies are being made so that the surface of the paper is somewhat spongy. The model 28 is capable of making the same number of carbon copies as present machines, and because the same hammer blow is applied for all characters, the printing impression is more uniform than that of a type-bar machine. This characteristic is very important when a large number of copies are being made. The printing blow can readily be increased for multiple copy work by adjusting the tension of a spring by means of a manually operated position lever, Fig. 4. Due to the simplicity of the printing mechanism, it is possible to make the shift from figures to letters and vice versa within the printing mechanism rather than by raising and lowering the platen roller. This simplifies paper handling and improves the readability of the printed record, since the paper remains stationary at all times except during line feed.

The ribbon spools are mounted on the machine itself rather than on the type carriage, thus providing a straight course for the ribbon travel. This not only facilitates changing ribbons, since the path is obvious and the number of guides is a minimum, but in combination with the more gentle blow of the new type hammer, it results in increased life of the ribbon as compared to the older machines.

A newly designed clutch is another machine element that improves operation, reduces maintenance, and contributes to good receiving margins; that is, ability to tolerate distortion of signal pulses. Clutches of the new design are used not only for driving the selector cams, but for the various other power actions, such as moving the typebox, feeding the paper, spacing, etc. This clutch is an all-steel internal-expansion friction clutch, that disengages in the stopped condition, whereas the older clutch depends upon slippage between felt washers and steel plates when a stop is interposed, so that the driven member is mechanically held from turning. Fig. 7 illustrates the method of operation. The continuously rotating driving member is a steel drum, the inner surface of which is grooved, hardened, and ground to give a flat surface on the tops of the grooves. The grooves between the flat surfaces permit wear products to fall away from the working surfaces. Within this drum, two hardened steel members act as drive shoes and are pressed into contact with the rotating drum by a spring-operated pry bar. The leverage system is so designed that through a system of very rigid force-multiplying levers, a small spring produces high normal pressure between the hardened steel friction surfaces. Since the clutch disengages in the idling position, the load on the motor at that time is very small. Life of the clutch equals that of the rest of the machine.

Performance of nonprinting operations, such as line feed, carriage return, and shift, is controlled by a new mechanism contained in a separate subassembly called a *stunt box*. This unit, which is accessible from the rear, extends across the full width of the typing unit, and engages code bars that also extend across the machine. The function bars, Fig. 8, of the stunt box engage notches in the code bars. The stunt box has 42 slots, each of which may hold a function bar capable of responding to an assigned code, making it possible to control 42 functions. Approximately 10 are reserved for the common functions, such as line feed, carriage return, and shift, and the remaining 32 are available for special purposes. Stunt boxes are interchangeable. The stunt box may be arranged to perform the functions of the sequential selector unit, shown in Fig. 9, which controls circuits from groups of character combinations sent in a predetermined sequence.

The keyboard mechanism in the model 28 is also different from earlier designs. When the operator depresses a key, a latch is tripped which permits the code bars to move endwise by spring action. In the older machines, depressing a key moved the code bars directly, so that there was considerable variation in the forces required for different code combinations. Thus the new action results in a lighter, shallower, and more uniform key touch.

When the key lever is depressed, one of the bars that moves longitudinally trips the clutch latch and allows the clutch to engage a cam-operated mechanical distributor. This causes the code pattern to be translated into a start-stop

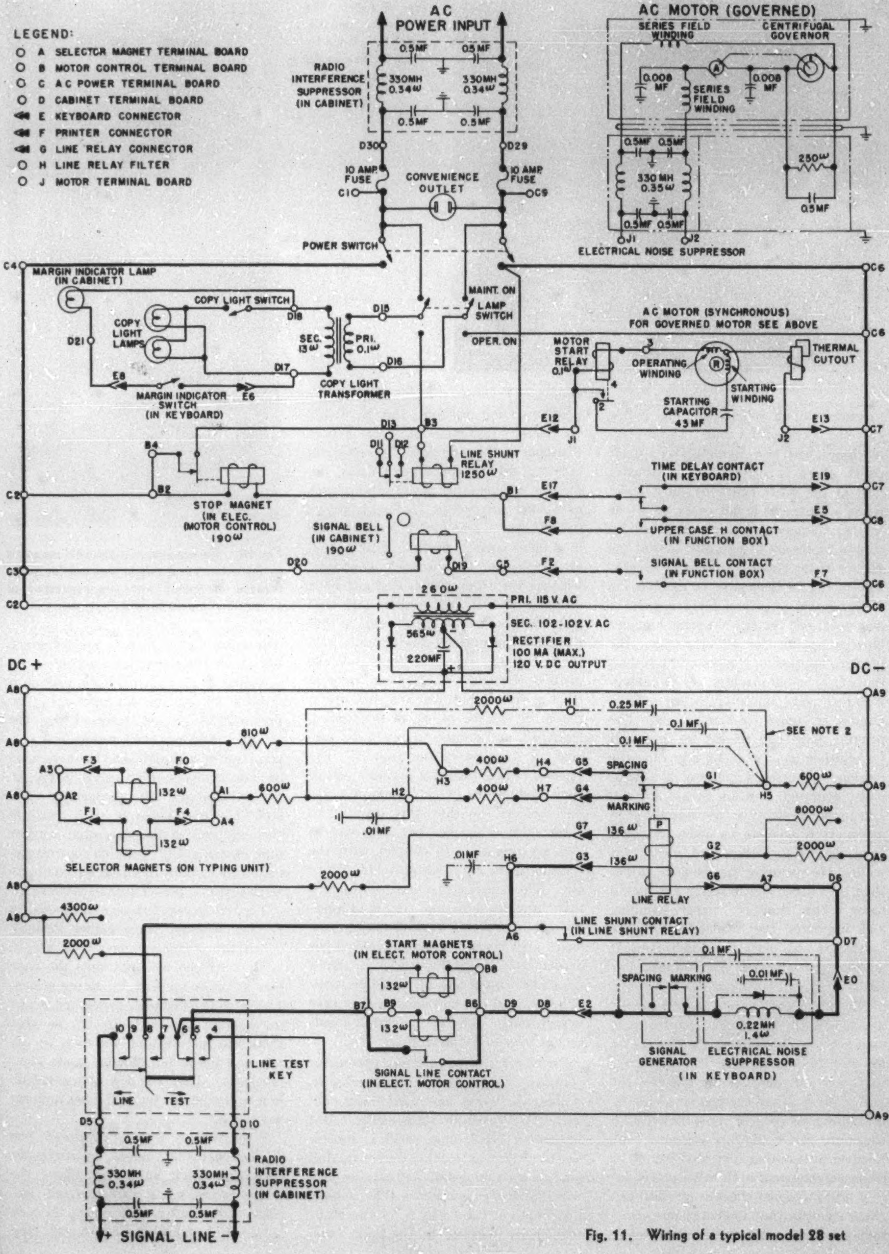

LEGEND:
○ A SELECTOR MAGNET TERMINAL BOARD
○ B MOTOR CONTROL TERMINAL BOARD
○ C AC POWER TERMINAL BOARD
○ D CABINET TERMINAL BOARD
E KEYBOARD CONNECTOR
F PRINTER CONNECTOR
G LINE RELAY CONNECTOR
○ H LINE RELAY FILTER
○ J MOTOR TERMINAL BOARD
AC POWER INPUT
RADIO INTERFERENCE SUPPRESSOR (IN CABINET)
0.5MF
330MH 0.34ω
D30
D29
10 AMP. FUSE
CONVENIENCE OUTLET
C1
C9
POWER SWITCH
AC MOTOR (GOVERNED)
SERIES FIELD WINDING
CENTRIFUGAL GOVERNOR
0.008 MF
A
330 MH 0.35ω
250ω
J1
J2
ELECTRICAL NOISE SUPPRESSOR
C4
MARGIN INDICATOR LAMP (IN CABINET)
COPY LIGHT SWITCH
COPY LIGHT LAMPS
D18
D15
D21
SEC. 13ω
PRI. 0.1ω
D16
D17
COPY LIGHT TRANSFORMER
E8
MARGIN INDICATOR SWITCH (IN KEYBOARD)
E6
MAINT. ON
LAMP SWITCH
OPER. ON
C6
AC MOTOR (SYNCHRONOUS) FOR GOVERNED MOTOR SEE ABOVE
MOTOR START RELAY 0.1ω
OPERATING WINDING
R
STARTING WINDING
THERMAL CUTOUT
STARTING CAPACITOR 43MF
E12
E13
C7
D13
B3
D11
D12
B4
LINE SHUNT RELAY 1250ω
TIME DELAY CONTACT (IN KEYBOARD)
E19
C2
B2
STOP MAGNET (IN ELEC. MOTOR CONTROL) 190ω
B1
E17
F8
UPPER CASE H CONTACT (IN FUNCTION BOX)
E5
C8
SIGNAL BELL (IN CABINET) 190ω
C3
D20
D19
G5
F2
SIGNAL BELL CONTACT (IN FUNCTION BOX)
F7
260ω
PRI. 115 V. AC
SEC. 102-102 V. AC
565ω
220MF
RECTIFIER 100 MA (MAX.) 120 V. DC OUTPUT
DC +
DC −
A8
A9
2000ω
H1
0.25 MF
0.1 MF
SEE NOTE 2
810ω
A3
F3
F0
132ω
A2
A1
600ω
H3
400ω
H4
G5
SPACING
H5
F1
F4
A4
H2
H7
G4
MARKING
G1
.01 MF
SELECTOR MAGNETS (ON TYPING UNIT)
G7
136ω
P
8000ω
H6
G3
G2
G6
A7
D6
LINE RELAY
4300ω
A6
LINE SHUNT CONTACT (IN LINE SHUNT RELAY)
D7
START MAGNETS (IN ELECT. MOTOR CONTROL)
B8
B7
B9
B6
D9
D8
E2
SPACING
MARKING
0.01 MF
E0
0.22 MH 1.4ω
10 9 8 7 6 5 4
LINE TEST KEY
LINE
TEST
SIGNAL LINE CONTACT (IN ELECT. MOTOR CONTROL)
SIGNAL GENERATOR
ELECTRICAL NOISE SUPPRESSOR
(IN KEYBOARD)
D5
D10
RADIO INTERFERENCE SUPPRESSOR (IN CABINET)
+ SIGNAL LINE −

Fig. 11. Wiring of a typical model 28 set

electric signal, the signal itself coming from the signal generator.

The signal generator is a single contact assembly mechanically operated by the distributor. This contact has the form of a transfer switch and therefore permits either open or closed signal transmission or transmission of signals of alternate polarity. The contact is mounted in a metal box for mechanical protection and shielding against radio interference, and requires no adjustment other than in positioning the box itself.

The cabinet for the new machine was designed to suppress machine noise, improve operating convenience, and provide better appearance. The equipment is housed in a new floor model, Fig. 1, with all mechanical controls brought to the front so that the machines can be mounted side by side in rows. Even the manual platen crank has been eliminated and replaced with a rapid motor-driven feedout controlled by a button on the keyboard. A lamp within the cabinet illuminates the copy, and the angle of the window above the copy has been chosen so that glare is practically eliminated. The upper section of the cabinet swings open to provide access for insertion of paper and ribbons and for maintenance. Fig. 10 illustrates how the equipment may be swung upward and forward to give access to both sides and rear of the machine.

Electrical Features and Circuits

Electric accessories such as the line relay, motor control relay, rectifier, fuses, etc., have been placed in a box behind the machine. The interconnections between the several units have been made simple and flexible, which permits installation of a standard machine where circuit termination requirements vary. In the cabinet below the printer there is a shelf on which a front panel is pivotally mounted, as shown in Fig. 10, providing mounting surfaces for auxiliary equipment if this is desired.

The equipment is driven by a synchronous motor when 110-volt regulated alternating current is available, or by a governed motor when unregulated alternating current or direct current is available. Printing telegraph equipment requires speed control of ±1 per cent. The governor used on the model 28, which is of new design will maintain the speed, and once adjusted, it will generally hold speed for the life of the motor brushes, even if the governor contacts wear or pit considerably. The older-type of governor requires frequent readjustment and maintenance and is subject to speed change due to comparatively slight wear or pitting of contacts.

The wiring of the model 28 Teletype set described herein, equipped with the most commonly used electrical features, is shown in Fig. 11.

Conclusion

Since operating experience indicates the model 28 page printer set requires less maintenance than other printing telegraph equipment, it is expected that its field of use will be extended to include more remote locations. The other units of the new line: the tape printer, perforator, and reperforator transmitter, have undergone extensive tests, which indicate they will give the same service as the page printer. These units also have new and novel features and will be available in the near future to serve the needs of printing telegraph users.

Reference

1. **Howard L. Krum.** *United States Patent No. 1,286,351*, **1919.**

No Discussion

Torque Requirements of a Radar Antenna

M. MARK
NONMEMBER AIEE

Synopsis: Results of wind tunnel tests on an experimental radar antenna are discussed. The torque required to rotate the antenna was seen to vary with the azimuth position, the elevation angle, the pivot location, the rotating speed, and the wind velocity. A correlation of the data based on dimensionless parameters was found.

A KNOWLEDGE of the loads imposed on a radar antenna at various wind velocities is important in order to obtain the optimum design for the driving and supporting mechanism. The torque loading in azimuth at various rotating speeds and wind velocities is a function of the position of the center of rotation or azimuth pivot location. Therefore, the engineer designing power drives for radar antennas would like to know the optimum position for the azimuth axis and the torque requirements for this pivot position. An experimental radar antenna was installed for testing in a wind tunnel in order to find the azimuth torque loadings associated with various pivot locations, antenna elevation angles, wind velocities, and rotating speeds. This paper discusses the general results of the testing. Elevation angle as used here will refer to the angle the antenna reflector screen makes with the vertical, positive when the concave face of the reflector points up, negative when down or depressed. Azimuth angle will refer to the angle measured in the direction of antenna rotation from the zero position, zero position being with the concave face of the antenna reflector directed into the wind.

Description of Equipment and Tests

The dimensions of the experimental reflector, a single-curvature parabolic type, were 26 by 84 inches. The reflector surface consisted of an aluminum screen, 3/16-inch standard expanded metal. The reflector was mounted on a special plate so that its location, and consequently the location of the pivot in azimuth, could be varied. The plate in turn was mounted on a pedestal incorporating a hydraulic drive, and rotated, turning the reflector. Fig. 1 shows the experimental antenna, installed in the wind tunnel, elevated to 45 degrees. A pressure-sensitive element in the supply line of the hydraulic drive provided a means for measuring

---

Paper 54-6, recommended by the AIEE Special Communications Applications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y. January 18–22, 1954. Manuscript submitted March 27, 1953; made available for printing October 5, 1953.

M. Mark is with the Raytheon Manufacturing Company, Newton, Mass.

torque. A calibration of the pressure was made by replacing the antenna reflector with a prony brake. Calibration curves were taken at various rotating speeds. For cases where autorotation (i.e., wind turning the antenna) occurred, no measure of the negative torque produced could be obtained, since the hydraulic pressure was not sensitive to this.

The tests could be divided into two groups. In the first, a search was made for the optimum pivot position (i.e., the pivot position in azimuth resulting in minimum peak torques). This was accomplished in a constant wind velocity of 60 miles per hour by varying the position of the reflector on the plate and comparing the torque readings. Rotating speeds of 8, 15, and 25 rpm, and four elevation angles, −13, 0, 22, and 45 degrees, were used. The second group of tests consisted of choosing an optimum pivot location and testing over a range of wind velocities from 40 to 100 miles per hour, with rotating speeds of 8, 15, and 25 rpm. Rotation was unidirectional.

Results

Typical of the results of the first group of tests are the curves for 15 rpm shown in Fig. 2, with the reflector at a −13-degree elevation angle. When the maximum torque occurring in a revolution was plotted against the pivot distance, which is the distance from the bottom center point of the reflector screen to the azimuth pivot measured in the plane of symmetry of the reflector, the results showed that the torque increase varied approximately parabolically with the distance from the optimum pivot. Optimum pivot positions were located for the four elevation angles. At any given elevation angle the optimum position was found to be the same, as closely as could be observed, for the three rotating speeds. With the reflector at 45 degrees elevation, the pivot distance for minimum torques was observed to be approximately 1/2 inch; for 22 degrees elevation the distance was 6 inches; at 0 degrees the value was 12 inches, and for the reflector depressed to −13 degrees the optimum distance was 16 inches.

The curves of Fig. 2 show interesting energy distributions. It can be seen that as the optimum pivot position (in this case 16 inches) is approached, the torque curve becomes more balanced for this screen type of reflector. For the case of the reflector depressed to −13 degrees in elevation and the pivot 7 inches from the bottom of the reflector screen, the torque curve is quite unbalanced, with large torques being required over the last half of the cycle and little or none over the first half. There are two major peaks, one in the region between 180 and 270 degrees in azimuth, the other near 360 degrees. As the pivot point is moved further away from the reflector, these peak torques are reduced, but more torque is required over the first half of the cycle. This process continues until the peaks become more or less balanced near the optimum pivot point (see curve for pivot distance of 15 inches in Fig. 2). As the pivot point is moved still further back, a large peak reappears in the torque curve, growing as the distance from the pivot increases. For a distance of 27 inches between pivot and reflector, it is seen that the peak torque occurs over the first half of the cycle in the region of 90 degrees in azimuth, and practically no torque is required over the last half. This is the reverse of what occurs on the other side of the optimum pivot point.

In all but two cases the pivot point was varied by changing the horizontal distance from the bottom center point of the reflector screen to the pivot point, keeping the pivot symmetrically located with respect to the horizontal transverse axis connecting the tips of the reflector. For two cases the pivot was displaced or offset from the symmetrical position and the effect on the torque curve noted. It was observed that this changed the location of the peak torque. With no displacement perpendicular to the plane of symmetry of the reflector, the peak torque for the optimum pivot occurred just after 180 degrees in azimuth, a second lower peak occurring in the region before 360 or 0 degrees. By displacing the pivot 1½ inches, the peak between 180 and 270 degrees was reduced, at the expense, however, of increasing the peak near 0 degrees. At a 3-inch offset this condition was even more pronounced. Consequently, it was seen that the difference between the peak torques at azimuth angles in the region of 0 and 180 degrees could be reduced by displacing the pivot from the plane of symmetry of the reflector.

Typical of the results of testing at a fixed pivot near the optimum and at wind speeds from 40 to 100 miles per hour are the curves shown in Fig. 3 for a 45-degree elevation angle.

Analysis of Data

The problem of predicting analytically the torque required to turn an antenna at any rotational speed in a given wind velocity is a difficult one. Configuration introduces a high degree of complexity. With antenna torque data available, empirical expressions offer the quickest and most direct solutions to the problem.

The wind tunnel data were correlated on the basis of dimensionless parameters. Consider a family of geometrically similar reflectors with given pivot position and

Fig. 1. Photograph of experimental antenna in wind tunnel (Wright Brothers Wind Tunnel, Massachusetts Institute of Technology)

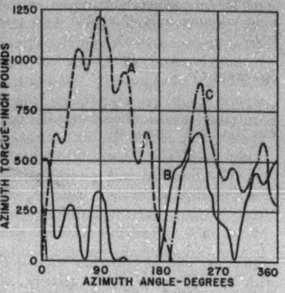


Fig. 2. Torque curves for antenna at −13 degrees elevation, rotating at 15 rpm in a wind velocity of 60 miles per hour

A. Pivot distance = 27 inches
B. Pivot distance = 15 inches
C. Pivot distance = 7 inches

elevation angle. Their size can be characterized by the reflector length or span D and the height H. The other parameters are wind velocity V, rotating speed N, and the density of air ρ (neglecting viscosity and compressibility). Generally, the peak torque occurring during a cycle T_{max} is of interest. Applying dimensional homogeneity,[1] two dimensionless parameters may be chosen and the relationship written

$$\left(\frac{T_{max}}{\frac{1}{2}\rho V^2D^2H}\right)=f\left(\frac{ND}{V}\right) \qquad (1)$$

Let a torque coefficient C_T be defined by

$$C_T=\frac{T_{max}}{\frac{1}{2}\rho V^2D^2H} \qquad (2)$$

Then equation 1 may be rewritten as

$$T_{max}=C_T\left(\frac{1}{2}\rho V^2D^2H\right) \qquad (3)$$

and

$$C_T=f\left(\frac{ND}{V}\right) \qquad (4)$$

The dimensionless torque coefficient is consequently a function of the rotating speed, wind velocity, and reflector span for a given reflector surface and configuration (i.e., shape, elevation angle, pivot location, back structure, and horn feed).

The data from the tests at an optimum pivot location described here were plotted in the form of equation 1 and the torque coefficient for this experimental antenna determined. The torque coefficient for a pivot in azimuth near its optimum location was seen to vary parabolically with the parameter $(ND)/V$. The results could be written

$$C_T=K_1+K_2\left(\frac{ND}{V}\right)^2 \qquad (5)$$

the dimensionless numbers K_1 and K_2 being functions of the configuration. For 0 degrees elevation and the pivot 11.3 inches from the screen, K_1 was found to be 0.037. K_1 was the same at −13 degrees elevation and a pivot distance of 14.8 inches. For 22 degrees elevation and the pivot 6 inches from the screen, K_1 was equal to 0.023; for 45 degrees elevation and a pivot distance of 1.67 inches, K_1 was 0.028. For all elevations and the pivot distances mentioned, K_2 was found to have the constant value of 33. The values for the pivot distances given were near their optimum. Consistent units were used in the calculations so that the two parameters of equation 1 were dimensionless. The scatter of the data about the faired parabolic curves was less than 10 per cent.

It might be expected that for a reflector surface, shape, and configuration different from that tested here, equations 3 and 4 still apply, but with the functional relationship of equation 4 to be determined. Consider a change in reflector surface only. A slatted type of reflector surface would probably give results similar to that for the screen surface tested here, for the case where the slats are at 0 degrees angle of attack with respect to the wind and the slat spacing is comparable to the screen opening. As the angle of attack of the slats increases, the surface could be expected to approach the solid type. For a solid type of reflector, the aerodynamic effects may be far more pronounced. For example, let the solid reflector be at an angle near 45 degrees in azimuth. It then is acting in a manner similar to an airfoil, deflecting the air flowing along its surface, which would result in larger pressure forces than for a screen or slatted type. (In addition, turbulence is present to a much greater extent.) This would mean that C_T should be much greater for the solid reflector.

The effect of configuration on the torque coefficient is, of course, difficult to predict. There are many variables. The back structure certainly influences the pressure forces, tending to make any open reflector surface approach the solid type. The horn feed acts like a fin, affecting the torques in the azimuth regions of 90 and 270 degrees. The optimum pivot location depends upon the configuration. The farther from the optimum pivot position, the higher the torque coefficient. It was seen that the optimum pivot location varied with the elevation angle for the experimental antenna tested. As the reflector was elevated and approached a more symmetrical body, the optimum pivot location moved closer to the reflector surface. In time, as more data becomes available, the precise effects of these many variables may become apparent and more accurate generalizations made. The considerations presented in this paper are offered as an addition to the fund of experience for future design work.

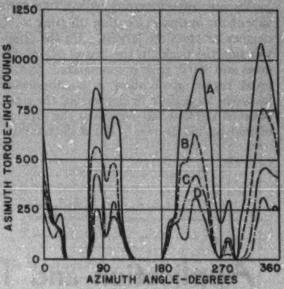


Fig. 3. Torque curves for antenna at 45 degrees elevation, rotating at 15 rpm with azimuth pivot 1.67 inches from reflector

A. Wind velocity = 100 miles per hour
B. Wind velocity = 80 miles per hour
C. Wind velocity = 60 miles per hour
D. Wind velocity = 40 miles per hour

Conclusions

1. The optimum pivot point (or center of rotation) of a radar antenna for azimuth torques gives more uniform torque requirements besides giving minimum peak torques. The further from the optimum, the more unbalanced becomes the torque curve, plotting torque versus azimuth angle (i.e., large torque required over half the cycle, little or none over the other half), and the larger the peaks.

2. As the antenna elevates and presents a more symmetrical body, the optimum pivot location moves closer to the reflector.

3. The increase in peak torque over that at the optimum pivot location varies approximately parabolically with the distance from the optimum point.

4. The antenna torque curves for a screen type reflector surface with given back structure and horn feed has major peaks in three azimuth regions—near 90 degrees, a little after 180 degrees (often between 180 and 270 degrees), and in the region before 360 degrees. Which of these peaks becomes the largest depends upon the pivot location. When the pivot is near optimum, the torque curve is well balanced and consequently the

peaks are approximately equal. On one side of the optimum location the peak at 90 degrees becomes the largest. On the opposite side the peaks in the region of 180 degrees and 360 degrees predominate. In this latter region pivot offset is an important factor in increasing or decreasing the peak between 180 and 270 degrees at the expense of that at 360 degrees. For a slatted-type reflector surface with slats at 0 degrees angle of attack the results should be similar. For a solid reflector surface, the peak torque will probably be in the azimuth region near either 45 degrees or 270 degrees.

5. Equation 3 gives the general type of relationship found to exist between maximum torque, wind velocity, and rotating speed for an experimental radar antenna in wind velocities from 40 to 100 miles per hour.

Reference

1. Advanced Dynamics (book), **S. Timoshenko, D. H. Young.** McGraw-Hill Book Company, Inc., New York, N. Y., 1948, p. 381.

No Discussion

Application and Transmission Features of a New 12-Channel Open-Wire Carrier System

K. E. APPERT
NONMEMBER AIEE

R. S. CARUTHERS
MEMBER AIEE

W. S. CHASKIN
NONMEMBER AIEE

Synopsis: A new, miniaturized, 12-channel open-wire carrier system, the Lenkurt *45A* system, is described. For use in the 40- to 150-kc frequency region, it co-ordinates fully in level and frequency allocations with present Lenkurt *42C* and Western Electric *J* carrier systems using this frequency space. Because of the inclusion of 2-pilot regulation for flat and slope gain control, supplemented by channel regulation with a range of about ±7 decibels (db), more stable circuit operation is obtained than in previous open-wire carrier systems. The design which results is much reduced in cost as compared to past systems in this frequency range, and affords an opportunity for considerable expansion in use of open-wire plant above the 35-kc range. This use is further enhanced by optionally supplied compandors, permitting many instances of use of wire lines transposed only for 35-kc operation. An important feature of the system is terminal design and frequency allocation co-ordinated for open-wire, single cable, and radio use. This permits great flexibility in interconnecting among such systems without need for intervening terminals. Miniaturization carried to a high degree results in a terminal occupying but $31^1/_2$ inches of vertical space on a telephone-type relay rack. Back-to-back operation permits as many as 96 channels in two 11-foot 6-inch bays. Extensive use is made of standardized plug-in units further supplemented by plug-in band filters and signaling options, to fit a wide variety of application requirements.

Applications and Objectives

DURING the past 15 years there has been limited application of carrier systems on open-wire lines extending to frequencies as high as 150 kc. Any higher frequency use is generally impractical, either due to the high losses occurring during ice conditions on the wires or to the high noise from lightning and static in wet weather. Notable among systems making full use of this 150-kc top range of frequencies are the Lenkurt *42C* system and the Western Electric *J* carrier. These and other systems have been limited in their application, either because they were not economical for use at short distances or because at long distances they degraded circuit performance to an objectionable degree. Another limitation has been the high cost of making lines operable for these systems above 35 kc. Although major reductions are being made in cost of line transpositions for use of these higher frequencies, as yet there has been no radical increase in the usage of 150-kc carrier systems. Even less to be expected are any radical additions of new wire lines for such purposes.

The *45A* system described in this paper, one terminal of which is shown in Fig. 1, is a new 12-channel arrangement co-ordinating fully in frequency allocation with the *42C* and *J* carrier systems, as well as with carrier systems now in common use below 35 kc. Twelve channels are obtained between 40 and 150 kc. Because of the low cost of this system it is expected that a much greater usage will now become practical on lines transposed only for channels operating below 35 kc. Even more impetus should be given to such applications of *45A* because of the superiority of its performance, stemming from better channel stability through use of channel regulation and double pilot repeater control, and through use of frequency shift signaling. Transmission-wise, no limitations exist in the use of the *45A* system for any part of a nation-wide toll switching network.

On most open-wire leads transposed for carrier systems below 35 kc, a number of pairs are found sufficiently free of ab-

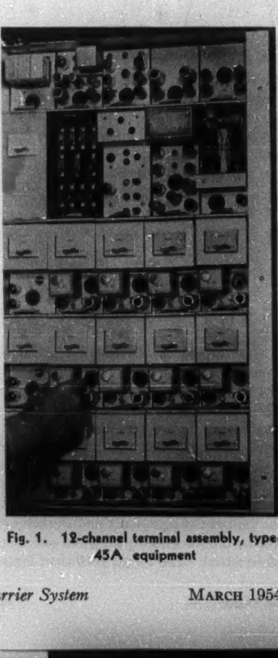


Fig. 1. 12-channel terminal assembly, type-45A equipment

Paper 54-58, recommended by the AIEE Wire Communications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 19, 1953; made available for printing November 18, 1953.

K. E. Appert, R. S. Caruthers, and W. S. Chaskin are all with the Lenkurt Electric Company, Inc., San Carlos, Calif.

The ideas incorporated in this new carrier equipment and the development of the system are the result of the co-operative effort of many persons in the Lenkurt organization. The success of the *45A* system is a tribute to their work. In acknowledgement, the authors wish to give credit—and the list of names is too long to include here—to all of the people who helped make this new system a reality.

sorption peaks to be usable to 150 kc. Frequently, coupling among such pairs would result in prohibitive crosstalk when more than one 150-kc system is installed on the lead. A compandor optionally supplied with the *45A* system will provide about 23 db of additional crosstalk suppression. This will permit utilization of many pairs of wire for 12-channel operation that are not now usable above 35 kc, with terminal and repeater equipment costs considerably below those of presently available equipment. This should result in a general elimination of the need for extensive wire line transposition work to 150 kc for a long time to come.

Some need still exists for long-haul expansion of 12-channel systems with 150-kc top range on leads already partially equipped. Again, *45A* will provide a superior and less expensive facility.

Another important characteristic of this new carrier system is that it provides an interconnectable type of equipment among 4- and 12-channel open-wire systems, cable carrier systems of the single-cable type, and radio systems without need for intervening terminals. Terminals are standardized for these various uses with channelizing, signaling, and most of the common carrier equipment alike. One system differs from another in a small number of group modulating and amplifier units used to obtain proper line frequency allocations. No modifications are required in open-wire, cable, or radio terminal equipment, whether it eventually terminates in the same or a different type of terminal facility.

Because this equipment must be used in a wide field of applications among independent and Bell System telephone companies, many options are required. Despite these demands, standardization is maintained at a high level with options provided by extensive use of plug-in parts. Plug arrangements are used in many ways. The channelizing equipment as it plugs into the equipment frame assembly is not identified as to channel frequency except by its frame position and insertion of a plug-in channel band filter in the channel unit. Likewise, East-West or West-East repeater assemblies as well as East and West terminal equipments are alike for East or West operations until identifying subassemblies are plugged in. Two-wire operation at terminals is obtained by plugging in a resistance hybrid; and optional 4-wire operation at the common telephone terminal levels of −13, +4, or −16, +7 is obtained through use of a particular plug-in assembly in place of the resistance hybrid. In-band or out-of-band signaling is also obtained from a plug-in option. Out-of-band signaling, provided through use of a frequency shift signal between 3,400 and 3,550 cycles, can be replaced by plug means, and in-band signaling using toll line dialing frequencies employed. In the latter case, an even wider speech band extending from 250 to 3,400 cycles can be obtained by plugging in a wider band filter in the channel equipment.

The *45A* system has been miniaturized to a very high degree. The terminal requires a vertical rack space of only $31^1/_2$ inches in a 19-inch telephone-type relay rack. Full use is made of 10-inch depth through the relay rack. This permits back-to-back relay rack arrangements, and as many as 96 channels can be mounted on two 11-foot 6-inch bays arranged in this manner. Extensive miniaturization of this system has been accomplished without problems of heat dissipation or any sacrifice in system performance or quality of its parts. Despite this high degree of miniaturization, the point has not been reached where further miniaturization has failed to reduce cost or improve performance.

A major step has been taken in providing inexpensive and simple means for checking and aligning channel equipment. Each plug-in unit is provided with an 11-pin vacuum-tube socket on its front face, allowing access to input and output speech and signal circuits. Without removing units from the bay, a cord plugged between the socket and a test assembly containing common types of telephone jacks allows tests to be made in normal telephone manner with associated measuring equipment. Only one such test assembly is needed for an entire office, and the jackfields normally found with such carrier equipment are not required for line-up or testing. The need for jacks to patch and reroute circuits is diminishing with increased dial operation. Provision of automatic channel regulation with its resulting high stability in the channel and signal circuit is a further factor in eliminating the need for jackfields and maintenance test positions.

Probably the most novel additional features in the *45A* system are the following:

1. The ability to adjust circuit net loss without removing the channel from service by measuring the level of one ever-present frequency-shift signal tone. This tone is also used for automatic channel regulation.
2. The use of a channel band filter of unusual design, whereby 4-kc spaced channels can be obtained that are not only in the broad-band channel bank category but in addition are wide enough to contain a frequency-shift signaling channel. This filter is even more unusual in that it is designed on a symmetrical basis with bandwidth from 250 to 3,750 cycles. The carrier can be shifted 4 kc and used at either edge of the filter to obtain an upper or lower side-band channel. The basic modulating plan of the *45A* system makes use of this feature to ob-

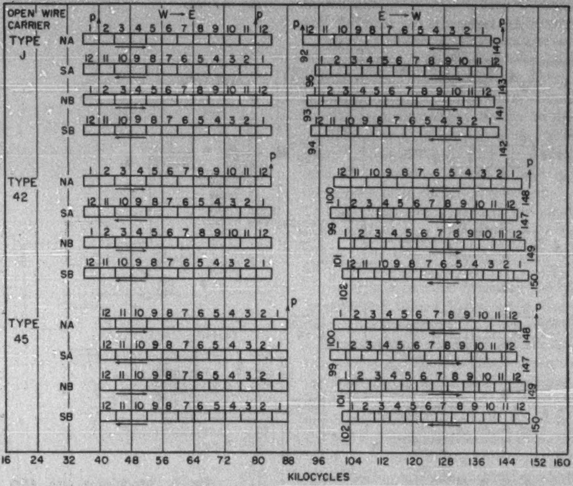


Fig. 2. Frequency allocation plan, 45A system, showing relation to type-J and -42C systems

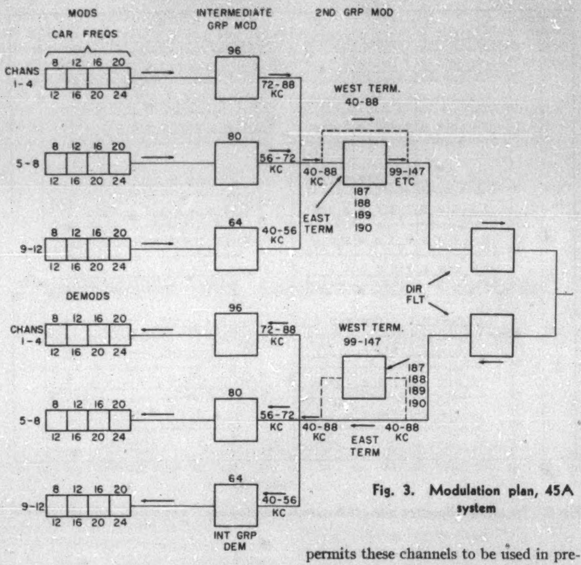


Fig. 3. Modulation plan, 45A system

tain upright or inverted frequency allocations on the line.

3. The use of carrier generating arrangements for the channel and group carrier supply frequencies using two crystal oscillators in the terminal to derive all of the frequency allocations of the various line spectra used. Each oscillator is held to close operating temperature limits in a thermostatically controlled oven. Frequency variation is of the order of one part per million even over long periods of time. Frequency divider methods are used to divide 96 kc into the eight frequencies required for channels and grouping equipment. A second oscillator at 187, 188, 189, or 190 kc sets the 12-channel group spectrum at the desired line frequency. No higher frequency is used in the system.

Frequency Allocation

In Fig. 2 the frequency allocations for the *45A* system, the Western Electric *J* carrier, and the Lenkurt *42C* system are shown. The *45A* system frequencies, though staggered and inverted as in the *J* and *42C* systems, differ in using channel numbers in identical order on the line, regardless of whether *NA*, *NB*, *SA*, or *SB* systems are used. Inversion is obtained by inverting the channels rather than the groups. In addition to the simplification and uniformity that results, a further benefit comes from the fact that the first four channels applied to the line are at the lowest frequency positions regardless of the staggered allocation chosen. This permits these channels to be used in preferred classes from noise, crosstalk, and attenuation (as in sleet) standpoints. Equipment is packaged so that it can be supplied in 4-channel units if desired.

In the low group, the *45A* system line frequency allocation is 4 kc higher than in the *J* and *42C* systems. This permits joint usage not only with the Bell system *C* carrier but also with the Lenkurt *33A* type carrier with line frequencies as high as 35 kc. Advantage can thus be taken of the *33A* system method of cheaply adding or dropping single channels on a way-station basis along an open-wire lead while allowing usage of the larger 4- to 12-channel *45A* groups for longer distances.

Basic Modulating Plan

The *45A* system makes use of basic 4-channel groups in the frequency range of 8 to 24 kc. These are modulated in three group modulators to form a 12-channel band at 40 to 88 kc. This group is either used directly on the line to form the low group for West to East transmission, or passed through an additional modulating stage with a carrier at 187, 188, 189, or 190 kc (for *NB*, *NA*, *SA*, or *SB* allocations) to form the high group for East to West transmission. Thus, the low group allocation is obtained with only two stages of modulation; and the high group, with one additional modulator stage, uses no modulating frequency higher than 190 kc.

In Fig. 3 the basic plan of *45A* modulation is shown. Three sets of four transmitted channels and three sets of four received channels are shown at the left. Each set of four channels uses the same carrier frequencies. These are 8, 12, 16, or 20 kc to obtain upright side bands, or 12, 16, 20, and 24 kc to obtain inverted side bands. The transmitted channels and the received channels must be inverted independently to meet the needs of the various line-frequency allocations. Thus, the 12 transmitted channels may be supplied with carriers at 8, 12, 16, and 20 kc, while the 12 received channels are simultaneously supplied with either 8-, 12-, 16-, and 20-kc carriers or 12-, 16-, 20-, and 24-kc carriers. Identical channel band filters are used in modulators and demodulators.

Each group of four channels is modulated or demodulated between the 8- to 24-kc and the 40- to 88-kc band by means of modulators using carriers at 96, 80, and 64 kc. Group modulator and demodulator carrier frequencies are the same regardless of the line allocations. Fig. 3 shows at the right the second group modulator used in the transmitting circuit at East terminals and in the receiving branch at West terminals.

System Description

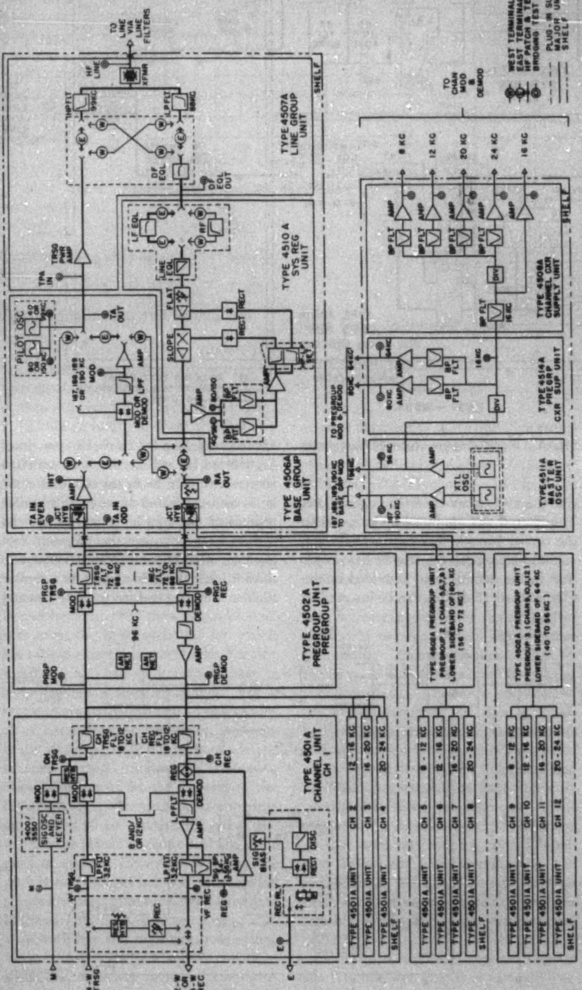
Terminals

A block diagram of a typical system terminal with built-in signaling is shown in Fig. 4. Voice-frequency signals enter the channel unit at a plug-in voice-frequency terminating subassembly. Different versions are provided for either 2-wire or 4-wire operation. The oscillator in the signaling circuit generates tones of 3,400 and 3,550 cycles per second under control of incoming signals on the *M* lead.

On channels not requiring out-of-band signaling, the plug-in signaling subassembly is replaced with a subassembly which makes the connections necessary to transmit a steady, regulating tone of 3,550 cycles. As only one out-of-band tone is transmitted, additional frequency spectrum is available for the speech circuit. The channel bandwidth can be increased by replacing the 3.2-kc low-pass filters with filters having a passband extending to about 3.4 kc.

The system pilot frequencies are generated by crystal oscillators which are contained in a plug-in subassembly of the base group unit. The subassembly for use at the West terminal includes 40-kc and 80-kc pilot oscillators. Internal circuits of this subassembly connect the base group modulator to the regulator output

Fig. 4. Block diagram, type-45A terminal


TO LINE VIA LINE FILTERS
TYPE 4507A LINE GROUP UNIT
TYPE 4510 A SYS REG UNIT
TYPE 4506A BASE GROUP UNIT
TYPE 4502A PREGROUP UNIT PREGROUP 1
TYPE 4501A CHANNEL UNIT CH 1
TYPE 4508A CHANNEL CXR SUPPLY UNIT
TYPE 4514A PREGRP CXR SUP UNIT
TYPE 4511A MASTER OSC UNIT
TYPE 4502A PREGROUP UNIT PREGROUP 2 (CHAN 5,6,7,8) LOWER SIDEBAND OF 120 KC (56 TO 72 KC)
TYPE 4502A PREGROUP UNIT PREGROUP 3 (CHAN 9,10,11,12) LOWER SIDEBAND OF 64 KC (40 TO 56 KC)
TYPE 4501A UNIT
SHELF
WEST TERMINAL
EAST TERMINAL
HF PATCH & TEST POINT
BRIDGING TEST POINT
PLUG-IN SUB-ASSEMBLY
MAJOR UNIT
SHELF
TO CHAN MOD & DEMOD
TO PREGROUP MOD & DEMOD
8 KC
12 KC
20 KC
24 KC
16 KC
4-W TRSG
2-W OR 4-W REC
M
E

and the base group transmitting amplifier directly to the line group unit. The subassembly for use at the East terminal includes 99-kc and 150-kc oscillators plus internal circuits to connect the base group modulator into the transmitting circuit and the regulator output directly to the junction hybrid. Thus, use of any base group unit at either an East or a West terminal is possible by merely inserting the proper pilot oscillator subassembly. The line group unit is arranged for use at either an East or a West terminal by inserting the proper plug-in directional filter equalizer subassembly.

The signals received from the distant terminal are passed through the directional filter (low-pass at East terminals or high-pass at West terminals) and the system regulator. The two pilot frequencies are filtered from the band of signals present at the output of the regulator and are amplified, rectified, and used to control the flat and slope compensation given the band of channels by the regulator. The regulator, in conjunction with the various fixed equalizer networks, acts to equalize the attenuation of the line so that all channels are at about the same level at the regulator output. At East terminals, the low-frequency 12-channel group is passed directly from the regulator to the junction hybrid of the base group unit. At West terminals, the high-frequency 12-channel group is passed through a step of group modulation to position the group in the low-frequency band from 40 to 88 kc.

Repeaters

A basic *45A* system repeater consists of two line group units and two system regulator units arranged as shown in the functional block diagram of Fig. 5. Each repeater also includes power distribution and fusing facilities and a terminal strip for making external connections. Units of one repeater are plugged into a single shelf. The line group units and system regulator units are identical with those used at terminals.

Terminal Carrier Supply

All the channel, pregroup, and base group carrier frequencies required for terminal operation are supplied from one carrier supply system. Accuracies of about one part per million are obtained through use of two oven-controlled crystal oscillators. Channel and pregroup carrier frequencies are obtained by a frequency-division process.

Connections to the channel modulators and demodulators of the proper carrier frequencies are made by use of plug-in connectors. Operation of two terminals

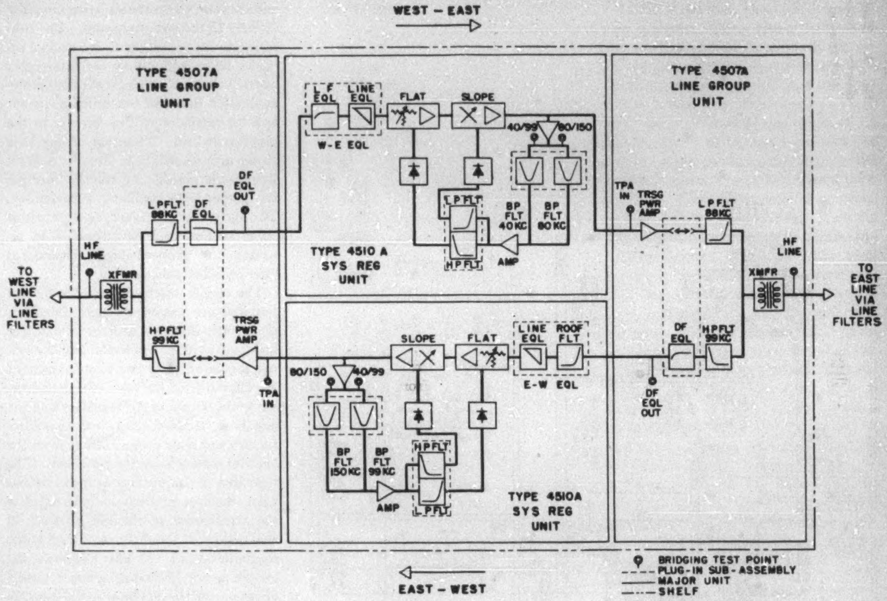

WEST – EAST
TYPE 4507A LINE GROUP UNIT
TYPE 4510 A SYS REG UNIT
TYPE 4507A LINE GROUP UNIT
L F EQL
LINE EQL
W-E EQL
FLAT
SLOPE
40/99
80/150
BP FLT 40 KC
BP FLT 80 KC
L P FLT
H P FLT
AMP
DF EQL OUT
L P FLT 88KC
DF EQL
HF LINE
XFMR
TO WEST LINE VIA LINE FILTERS
H P FLT 99KC
TRSG PWR AMP
TPA IN
TPA IN
TRSG PWR AMP
L P FLT 88KC
HF LINE
XMFR
TO EAST LINE VIA LINE FILTERS
DF EQL
H P FLT 99KC
DF EQL OUT
SLOPE
FLAT
LINE EQL
ROOF FLT
E-W EQL
80/150
40/99
BP FLT 150KC
BP FLT 99KC
AMP
H P FLT
L P FLT
TYPE 4510A SYS REG UNIT
BRIDGING TEST POINT
PLUG-IN SUB-ASSEMBLY
MAJOR UNIT
SHELF
EAST – WEST

Fig. 5. Block diagram, type-45A repeater

from a single carrier supply, to facilitate maintenance or during emergencies, can be accomplished by patching test cords between the carrier distribution circuits of the two terminals. A block diagram of the terminal carrier supply system is included in Fig. 4. The master oscillator unit is equipped with the two oven-controlled crystal oscillators. One of the oscillators generates the carrier used for base group modulation at East terminals and for base group demodulation at West terminals. Its frequency is 187, 188, 189, or 190 kc, depending upon the line allocation. The other oscillator generates a 96-kc carrier for pregroup unit no. 1. The 96-kc output is also supplied to the frequency division network of the pregroup carrier supply unit where frequencies of 80, 64, and 16 kc are derived. A frequency of 16 kc is passed to the channel carrier supply unit where another frequency division network derives the five frequencies required for the channel modulators and demodulators.

Fig. 6 shows the frequency division method whereby 96 kc is divided down to obtain channel and pregroup carrier frequencies. Ninety-six kc is fed to a modulator with 80 kc fed into the normal carrier branch. The difference modulation product of 16 kc is fed from the modulator to a nonlinear varistor device used to derive the fifth harmonic of 16 kc, or 80 kc. The filtered and amplified 80 kc is fed to the carrier branch of the aforementioned modulator circuit. Experience has shown that in the arrangement provided in the *45A*, at least 15-db loss can be inserted in the feedback loop before the division process fails to be self-starting. In addition to generating 16 kc in the modulator, 64 kc is also generated through third-order process of twice 80 kc minus 96 kc. In this manner not only the pregroup frequencies of 96, 80, and 64 are derived, but in addition 16 kc is developed for a similar divider unit to generate the five channel

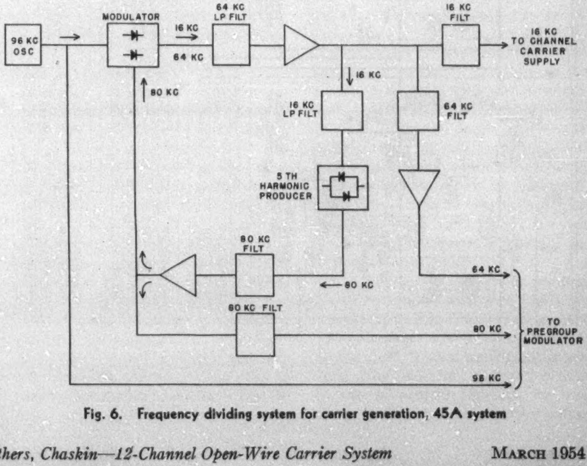

96 KC OSC
MODULATOR
16 KC
64 KC
80 KC
64 KC LP FILT
16 KC FILT
16 KC TO CHANNEL CARRIER SUPPLY
16 KC
16 KC LP FILT
64 KC FILT
5 TH HARMONIC PRODUCER
80 KC FILT
80 KC
80 KC FILT
64 KC
80 KC
96 KC
TO PREGROUP MODULATOR

Fig. 6. Frequency dividing system for carrier generation, 45A system

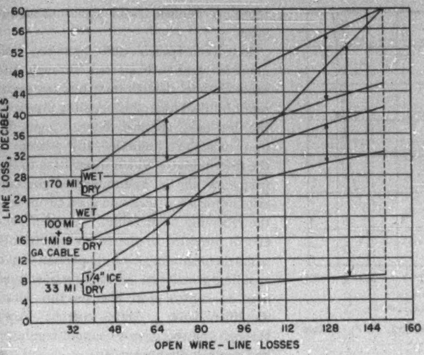


Fig. 7. Regulation characteristics, 45A system

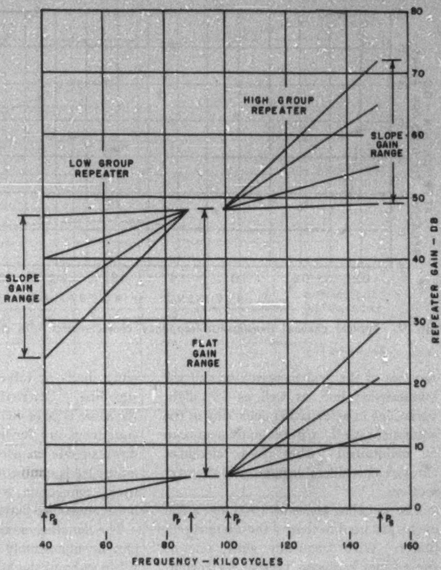


Fig. 8 (right). Low and high group gain characteristics, type-45A repeater

frequencies in the channel carrier supply unit.

Slope Equalization and Regulation

A wide variety of line attenuation characteristics are encountered in use of systems of the *45A* type. Some examples of these line conditions are shown in Fig. 7. Although in dry and wet weather the line losses and slopes are quite predictable and complicated only by the amount of entrance cable involved, during ice conditions the line losses and slopes may vary over a wide range. In certain cases high losses may be accompanied by extremely steep slope conditions in the line attenuation characteristic, and in others the high loss may occur with slopes no greater than in wet weather. For the amount of sleet that must be coped with and for which the repeater spacings must be set, the dry weather losses are so low that there is little flat loss or slope remaining. For a repeater regulating system to encompass all of these varying line losses, it is virtually essential that a 2-pilot regulating scheme be used. In the *45A* system, a slope pilot is used at the top end of the high group at 150 kc with a corresponding flat pilot at 99 kc. In the low group, the flat pilot is placed at 80 kc and the slope pilot at 40 kc. Choice of 80 kc as the pilot is made to get below the absorption region frequently found on open-wire lines in the vicinity of 88 to 95 kc.

In both low and high group repeaters, the maximum gain that is set in the vicinity of the crossover region of the directional filters—that is, 88 to 99 kc—is determined by the maximum gain that can be inserted by the respective flat pilot regulators. The slope pilot controls at the extreme outer edges of low and high group bands, as they vary the amount of slope gain, have virtually no effect on the gain at the flat pilot frequencies. In Fig. 8 the gain characteristics of low and high group repeaters are shown. In the high group, dry weather conditions are met by the slope control turning slope gain to near-minimum value. In the low group, the slope control works in the opposite direction, increasing gain at the slope pilot frequency in moving toward dry weather gain setting. This added gain is offset by the increase in loss introduced by the flat pilot frequency. The basic operation of the system regulator in both terminals and repeaters is shown in the terminal and repeater block diagrams of Figs. 4 and 5. A system of feedback amplifiers and variable loss networks are used which are controlled by thermistor elements. At the output of the regulator, the two pilot frequencies are selected by crystal band-pass filters, and after amplification and rectification, the direct current controls the thermistor resistance and in turn, the loss of the flat and slope regulator amplifiers.

Although a line characteristic is shown for 170-mile wet line, it is virtually certain that this would never be a practical operating condition because of the high noise levels from static existing at this time. This line loss is made indicative of a combination of cable and wet weather losses that might be expected during periods of normal operation on circuits of extreme length. It is a basic part of the plan for using 2-pilot regulation systems of the type in the *45A* that repeater gain would not always be adequate for the amount of sleet losses encountered. In such cases it would be expected that the repeater gain would expand to about a 48-db maximum across the entire low group, and to a gain ranging from 48 to 72 db at 150 kc in the high group. Even though such gain did not compensate fully for the sleet line loss, it would be expected that subsequent repeaters and the terminal would make up as far as possible for any deficiencies in gain of preceding repeaters. As a further aid to this, the additional channel regulation margin of about 7 db existing at the terminals would be called on during extreme sleet conditions to attempt retention of circuit equivalents at satisfactory values.

Channel Regulation

To compensate for any transmission discrepancies not corrected by the system regulator, each channel unit is equipped with an automatic channel regulator which has a range of about ± 7 db. Individual channel regulation will correct for differences between the frequency charac-

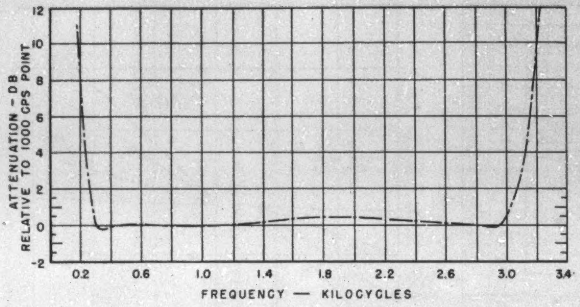


Fig. 9. Typical channel transmission-frequency characteristics with signaling, 45A system

teristics of the system regulator and the transmission line, as well as for slight variations in attenuation occurring in the equipment itself. Circuit equivalents can be maintained within close tolerances through a considerable number of repeater sections.

The signaling tones of 3,400 and 3,550 cycles are used to control the channel regulator. With frequency shift, one of these tones is always present for regulation.

The channel regulator uses a thermistor bridge between the demodulator and receiving channel band filters. The signaling frequency controls a direct current through the thermistor and varies its resistance. When the input to the channel falls below normal the thermistor bridge becomes more unbalanced, tending to bring the input to the demodulator back to normal.

Signaling Facilities

A system of built-in, out-of-band, frequency-shift signaling is usually furnished for each channel of the system. All-electronic varistor keying is used at the sending terminal, while a polar relay is operated at the receiving end. The circuits are terminated in *E*, *M*, and *F* leads which can be arranged to operate with dial and either half- or full-duplex d-c ringdown signaling. External converter units of the same type of construction as other *45* equipment are furnished when a-c ringdown signals are used. The channel signaling leads can be arranged for back-to-back connection without using pulse link connectors between *E* and *M* leads.

The signaling circuits are contained in a plug-in subassembly of the channel unit. When in-band single frequency dialing or some type of externally applied voice-frequency signaling is used, or in telegraph use when signaling is not required, this subassembly may be replaced with a channel regulator subassembly that makes connections for the transmission of a steady channel pilot of 3,550 cycles.

The signaling circuits are shown in the block diagram of the system terminal. External signals, impressed on the *M* lead, switch a capacitor in and out of the signaling oscillator frequency-determining circuit, causing it to generate alternately tones of 3,400 and 3,550 cycles. These tones are applied to the signaling modulator.

Received speech and signaling frequencies are both passed through the same channel demodulator and are separated by the band-pass signaling filter and the low-pass channel filter. Signaling frequencies are amplified and passed to the discriminator which is sensitive to the frequency shift. The output of the discriminator is rectified and used to operate the receiving polar relay. A potentiometer is provided in the receiving circuits to correct signaling bias. A portion of the signaling frequency power is rectified and used to control grid bias on the signal amplifier tube. Plate current flows through the thermistor channel regulator and controls channel gain.

The signaling circuits can be adapted for operation with all usual arrangements of d-c ringdown or dial signaling equipment using battery, ground, or open circuits at the switchboard. Strap connections in the signaling subassembly allow use of a variety of *M* lead signals and also make various connections to the *E* or receiving lead to provide battery, ground, or open as the polar relay is operated by the received frequency-shift signals.

Optional Compandors

The *45A* system is designed to give toll-quality performance with respect to equipment noise, distortion, and crosstalk products, without the use of compandors. Therefore, compandors are not built in but are available as a standard option to permit system application to lines which would normally be unsuitable because of noise or crosstalk.

Use of the compandor on a channel will provide an effective signal-to-noise

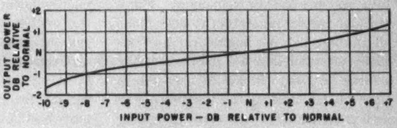


Fig. 11 (right). Typical regulator characteristics, type-45A channel

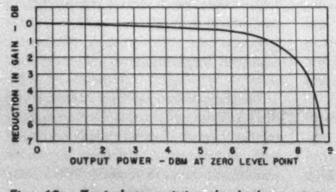


Fig. 10. Typical transmitting load characteristics, type-45A channels

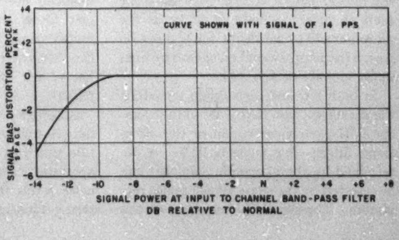


Fig. 12 (right). Signaling characteristics, 45A system

advantage of about 22 db. When equipped with compandors, two or more *45A* systems can often be operated on 35-kc transposed lines which have suitable attenuation characteristics. The noise advantage gained may also permit toll-quality transmission to be obtained from lines with high inherent noise levels.

System Performance

Channel Transmission Frequency Characteristics

Over-all frequency characteristics of a typical non-compandored channel of type-*45A* equipment with out-of-band frequency-shift signaling are shown in Fig. 9. The measurements were made with carrier terminals connected back-to-back and reflect no line irregularity contributions. Although characteristics will vary slightly from channel to channel, the curve shown is representative for conditions in which the system performance is determined principally by the carrier equipment. Generally, frequencies of about 250 and 3,100 cycles have a net loss 3 db greater than the 1,000 cycles net loss. Over the frequency band from about 300 to 3,000 cycles, any net loss variations from the 1,000 cycles net loss should not exceed about 3/4 db.

When the built-in, out-of-band signaling is not used, a low-pass channel filter is available to extend the upper 3-db point of the transmission frequency characteristics to about 3,400 cycles.

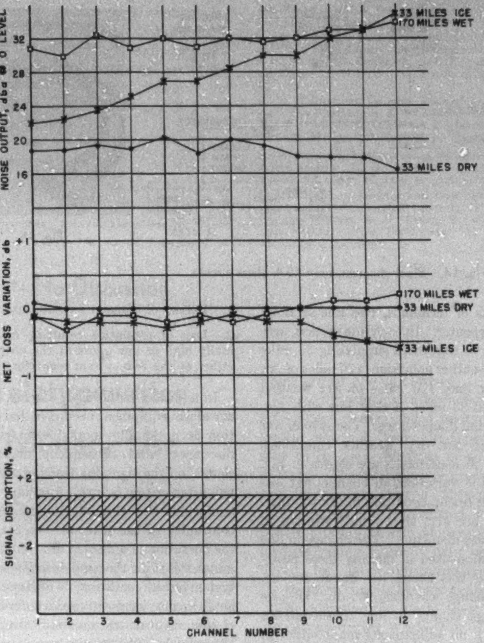


Fig. 13. Performance characteristics, 45A system

Operating Level and Load Characteristics

Two-Wire Drop Levels

Nominal 2-wire transmitting level at the drop side of the resistance hybrid is 0 db. The loss through the hybrid establishes the level at the input to the channel modulator at −16 db. A continuously variable pad is included in the receiving branch of the hybrid. Adequate gain is available so that receiving levels on the drop side of the hybrid as high as about +5 db can be obtained; however, actual maximum level permissible will depend on the resistance hybrid balance which is governed by the impedance of the connected drop circuit. The ideal condition is obtained with a drop circuit having a pure, resistive 600-ohm characteristic impedance.

Four-Wire Drop Levels

Nominal transmitting and receiving 4-wire voice-frequency levels of −16 and +7 db, or −13 and +4 db respectively, can be established by the use of the proper plug-in 4-wire voice-frequency terminating subassembly in the channel units. When external 4-wire terminating units are used, adjustments included in each branch of the terminating unit permit additional flexibility in 2-wire or 4-wire extensions.

The pregroup and common equipment have sufficient load capacity for all expected speech loads. To prevent severe overloading by any one channel output, each channel limits the peaks of very loud speech passed to the common circuits. Typical load characteristics of type-*45* channels are shown in Fig. 10. Speech is satisfactory for even the loudest of talkers affected with this limiting action, provided nominal transmitting drop levels are employed.

Channel Regulator Performance

A typical measured characteristic of the channel regulator is shown in Fig. 11. The channel regulator should maintain the channel receiving drop level within about ±1 db of normal for variations of the level at the channel receiving band-pass filter over a range of +5 to about −7 db.

Signaling Performance

The signaling tones are transmitted at −16 db below 1 milliwatt referred to the system zero reference point. With transmission of pulses at speeds between 8 and 14 pulses per second, signal bias distortion is expected to be less than 1 per cent in all conditions for which the over-all system regulation is effective. Voltage variations of the plate or filament supply voltage of ±10 per cent have little effect on signal bias distortion. The use of the frequency-shift method of transmission gives reliable operation with good freedom from the effects of interference and noise.

Typical signal bias versus received level characteristics of the signaling circuits are shown in Fig. 12. Data were obtained by varying the received signal level at the channel band-pass receiving filter.

Noise and Circuit Stability Under Varying Line Attenuation Conditions

In Fig. 13 data are shown of the noise levels at channel outputs measured on an experimental laboratory setup involving

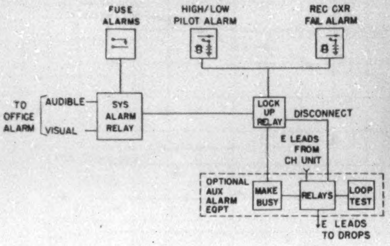


Fig. 14. Block diagram, type-45A alarm circuits

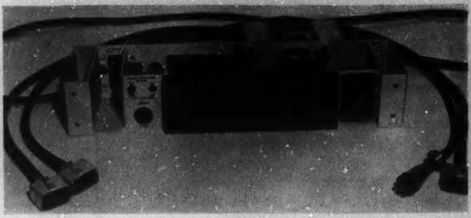


Fig. 15. Test and jack panel for use with 45A terminal

two sets of terminals, two line sections, and a repeater. In each line section, artificial lines were used simulating 33 miles of dry weather condition, 170 miles of dry weather line, 170 miles of wet weather line, and 33 miles of sleet line with 1/4-inch radial diameter ice. The system was lined up for dry weather conditions. Similar measurements are shown in Fig. 13 for the variations in channel net loss from the lineup condition in dry weather. Curves are also shown for variation in signal bias distortion. The change in line attenuation used in making these measurements was about 104 db at 150 kc. Slope change was about 45 db. Despite this large change in line characteristics and with no adjustments made either in the speech circuits or signal circuits, signaling circuit and channel net loss changes were virtually unmeasurable. The noise during sleet and wet weather conditions increased somewhat as planned for in the design. A careful disposition of amplifier gains and thermistor flat and slope regulator controls in the design of the repeater regulator keeps the repeater and terminal circuits free from overload and noise limitations throughout all expected line loss conditions.

Alarm Facilities

A simplified block diagram of the *45A* basic and auxiliary alarm equipment is shown in Fig. 14. Alarm facilities provide visual indication by lamps on the alarm units as well as contacts for connection to external visual and audible alarm circuits. Basic alarm equipment provided at terminals is operated by the following conditions:

1. Failure of any filament or plate fuse.
2. Loss of received carrier energy.
3. Loss of regulation resulting in abnormally high or low pilots at the output of system regulators.

At repeaters, alarm equipment is provided which is operated by the following conditions:

1. Failure of filament or plate fuse.
2. Loss of regulation resulting in abnormally high or low pilots at the outputs of either of the two system regulators.

In addition to the basic alarms, auxiliary alarm equipment needed in dial operation is optionally available to provide disconnect and subsequent make-busy signals on the signaling leads of all channels after system failure. Auxiliary alarm equipment can also be used to test the continuity of the transmission path and the operation of a distant unattended terminal, following the supposed clearing of a system failure condition. The basic alarm facilities provide positive alarm for failure of equipment or transmission common to 12 channels. The circuits associated with the received carrier alarm monitor the total carrier energy received at terminals. An alarm is indicated when the total carrier energy received falls below that received when speech is absent and only signaling tones are transmitted. Sufficient margin is provided so that noise will not operate the alarm.

Received carrier alarm and regulation alarm will also interrupt all transmission from the terminal. This will cause received carrier alarm at the distant terminal and regulator alarm at all repeaters. Thus, in the event of failure of equipment which causes system failure, alarm is provided at all repeaters and terminals. This is mandatory because of the interconnection between system failure alarms and required subscriber disconnect signals at the two terminals.

Test and Maintenance

It is not expected that telephone-type jackfields would normally be used with type-*45* terminal equipment. Instead, a panel shown in Fig. 15 is used providing jack access for all normal terminating test procedures. The test set is connected to the carrier equipment by using special test cords to patch from a plug or socket on the test set to test sockets on the equipment being tested. One such test set could be used for an entire office, in contrast to the large arrays of jackfields which were required with earlier carrier-systems.

An optional telephone set circuit may be added to provide monitoring, talking, and signaling facilities for any of the *45* system channels. For test and lineup, channel units need not be removed from their plugged-in positions. Instead, the resistance hybrid or 4-wire plug is removed from the channel unit and plugged into the socket provided on the test panel. A patch cord is then extended from the vacated socket on the channel unit to the test set. Standard telephone-type jacks on the test set are thereby placed in the voice-frequency and signal circuit giving access to 4-wire equipment and drop circuits, as well as *E* and *M* leads.

In addition to the use of the auxiliary test jack panel, bridging tests can be made using a vacuum-tube voltmeter inserted in an 11-pin vacuum-tube socket jack on each equipment unit. For further tests any unit can be removed from the frame position and restored to operating conditions through a cord with 20-pin plug and ack at the two ends.

Conclusion

The system described in this paper has undergone extensive development tests for nearly 1 year, and at present a considerable number of these systems are being put in operation throughout the operating telephone plants. Performance has met the objectives in a highly satisfactory manner. The inclusion of automatic channel regulation combined with frequency-shift signaling has proved to be an invaluable asset to system operation. Although usage through several winters of sleet storm exposure is required to prove fully the adequacy of the regulating arrangements provided, there is no doubt that the ranges afforded will cope with most line situations where wire lines remain intact.

The almost wholesale shift to the use of plug-in units and subassemblies has been an invaluable aid in system development, in factory production, and in field maintenance. With the completion of the channelizing group, and carrier supply equipment, a major part of the development was accomplished for new cable and radio systems. Development of only a few new group units and slight revisions in frame wiring are needed to place these systems in operation. Further extensions of the development are being aimed at utilization of the basic equipment in supergroup assemblies of 60 channels for radio and coaxial cable use of many channels. The basic terminal equipment will also be utilized in 4-channel group assemblies with frequency - frogging repeaters both above and below 35 kc on open-wire lines.

References

1. A New Telephone Carrier System for Medium Haul Circuits, **R. S. Caruthers, H. R. Huntley, W. E. Kahl, Ludwig Pedersen.** *Electrical Engineering*, vol. 70, Aug. 1951, pp. 692–97.

2. Type-O Carrier Telephone, **J. A. Coy, E. K. Van Tassel.** *AIEE Transactions*, vol. 71, pt. I, 1952 (Jan. 1953 section), pp. 428–37.

3. A Miniature Compandor for General Use in Wire and Radio Communication Systems, **F. S. Boxall, R. S. Caruthers.** *AIEE Transactions*, vol. 72, pt. I, 1953 (Jan. 1954 section), pp. 804–11.

4. A 12-Channel Carrier Telephone System for Open Wire Lines, **B. W. Kendall, H. A. Affel.** *AIEE Transactions* (*Electrical Engineering*), vol. 58, July 1939, pp. 351–60.

5. A Simplified 48-Channel Carrier Telephone System, **L. G. Erickson.** *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1493–1500.

6. Mechanical Aspects and Component Features of a New Twelve-Channel Open-Wire Carrier System, **A. G. Ewing, F. W. Frazee, Dale Welling.** *AIEE Transactions*, vol. 73, pt. I, March 1954, pp. 75–81.

No Discussion

A Fully Automatic Teletypewriter Switching Center for Military Use

LEITH JOHNSTON
NONMEMBER AIEE

R. C. STILES
ASSOCIATE MEMBER AIEE

THE United States Army Signal Corps has recently added another chapter to the colorful history of military communications. On Dec. 17, 1952 at Fifth Army Headquarters, Chicago, Ill., the Signal Corps placed in service its new teletypewriter switching center which automatically switches messages through the Army's vast world-wide teletypewriter network. Sufficient time has passed since the system was cut into service to provide adequate demonstration that the system is an important link in modern military communications. Behind the drama of the cutover to the new system lies a story of technical progress which has kept the Signal Corps abreast of the most modern developments in communications technique. Some of the many technical and operational problems encountered in the development of this highly complex system will be told here.

The story of the teletypewriter art is one which parallels to a certain extent the development of the telephone art. The first attempts at teletypewriter switching centers saw the manual interconnection of lines by an operator for point-to-point communication just as in the early days of telephone communication. As the art swung toward the transmission of separate messages, more elaborate switching centers were employed to route messages toward their ultimate destinations. The first commercial automatic switching centers were installed just prior to World War II. These centers were capable of receiving messages and automatically routing them toward their several destinations without the assistance of operators. Thus a teletypewriter switching center is similar in purpose to an automatic dial telephone system. One such system for commercial use has been described by W. M. Bacon and G. A. Locke.[1]

During World War II, teletypewriter communication assumed new importance. It was during this period that the military commands began to use large scale networks of teletypewriter channels and equipment to facilitate their task of extending communications on a world-wide scale. The networks which were formed for the military commands consisted of both Government-owned radio facilities, and lines or channels which were leased from commercial companies on either a full-time or a part-time basis and which were terminated in teletypewriter equipment which was largely owned by the Government and operated and maintained by personnel of the various commands. To minimize the number of lines required, each network contained a number of switching centers which functioned in these networks in much the same manner as a telephone central office functions in a telephone system. Each teletypewriter station was connected to a switching center either by a direct line or in some cases by means of a party line which served several stations.

One of the principal differences between a telephone system and a teletypewriter system, however, lies in the type of message to be conveyed from one point to another. The telephone requires point-to-point contact, but a teletypewriter message may be transmitted from its point of origin to one of several intermediate points where it may be reproduced for later retransmission when the required circuits become available.

The equipment in use at these switching centers during World War II was of the manual type, which required that a message into a switching center be reproduced on a printed and perforated paper tape. An operator was required to read the address code of a message and, following complete receipt of the message, was required to place the message tape temporarily in a storage rack. Subsequently an operator was also required to withdraw the message from storage and insert it into a transmitter as the appropriate line became available. This manual process frequently caused delays and also created a complex problem of personnel handling.

Following World War II, the Signal Corps began a search for newer and more efficient switching center equipment to put their networks on a permanent basis. The Signal Corps recognized that the task of switching a message to temporary storage and subsequent retransmission could be accomplished automatically by the

---

Paper 54-122, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 21, 1953; made available for printing December 2, 1953.

LEITH JOHNSTON and R. C. STILES are with the Automatic Electric Company, Chicago, Ill.

The authors gratefully acknowledge the assistance of the Signal Officer, Fifth Army Headquarters, for assistance rendered in the preparation of illustrations and diagrams for this paper.

Fig. 1. General view of switching center in operation at Fifth Army Headquarters

use of techniques commonly used in the art of automatic telephony. An examination of existing commercial systems found those systems to be inadequate for Army requirements. The Signal Corps prepared a specification covering a switching center which would meet Army requirements. The system described in this paper was developed to meet the provisions of that specification.

Characteristics of Military Messages

Before considering the characteristics of the equipment required to switch military messages the reader should have an adequate understanding of the types of messages involved in a military network. Military messages range in length from very short messages such as the average telegram message to long messages involving requisitions, reports, troop movement data, logistics reports, and even public relations information which is to be given to the press. Interspersed between these longer messages are also the extremely short ones dealing with the operation of the network itself. Messages of this type usually use three letter operating signals to indicate a particular situation, such as, for example, the use of the operating signal *ZDK* to instruct a switching center operator to rerun a particular message for checking purposes. This wide range of message lengths is an important factor in the design of equipment. The average military message, however, is considerably longer than the average commercial message, thus accounting partly for the need of special equipment.

Military networks are used by nearly all types of military installations. It is not likely that front line or even rear area troops deployed in a tactical situation would be using the facilities of a teletypewriter network which would use automatic switching. Behind these tactical troops, however, are the many administrative and supply organizations which must keep in constant communication with the zone of the interior. For these communications, the written word and speed of service are highly important. Within the zone of the interior, however, many types of military activities such as training commands; army posts, camps, and stations; recruiting centers; and supply points and depots use the networks to good advantage. Most of these stations are equipped with network facilities only when the volume of traffic to or from the station warrants the leasing of line facilities, thus insuring that most of the lines into a switching center will carry a reasonably heavy load of message traffic.

One of the prime considerations in the network is the highest degree of reliability, and it is primarily this factor which determines network procedure, message format, and equipment characteristics. Seldom in a commercial network does life or death to many men or the success or failure of a major undertaking such as a military operation depend upon specific messages being delivered to the proper addressee. This same consideration also is closely coupled to the need for particular messages being given priority over others. Thus in a military network the relative degree of precedence assigned to a message by its originator must be recognized and adhered to in the switching centers. Messages of extremely high precedence actually require interruption of lower precedence to clear the way for the high-precedence message. This particular characteristic of a military network is of utmost importance in the design of automatic-switching-center equipment.

In a network which operates on a worldwide basis it is most convenient to use a plan of station-call-letter assignments in which the call-letter assignment itself identifies the activity and its approximate geographical location. The first consideration is the identification of the network, since in modern military operation all services which operate their own networks must also be able to transmit messages of all other services. Joint agreement, therefore, has assigned the first letter of a station call sign, known as a routing indicator, to identify the service network involved. Since all of our military networks are composed only of switching centers each of which serves local nearby stations, the next point of consideration is the geographical location of the switching centers. For this purpose the world has been divided into a number of areas to each of which a definite letter has been assigned. Thus the second letter of a routing indicator serves to identify a particular area of the world. The third point of consideration is the particular switching center within the designated area and network. Thus the first three letters of the routing indicator identify a particular switching center. Succeeding letters of a routing indicator serve to identify a particular tributary station of the designated relay center. The routing indicator assignment plan is therefore of utmost importance to the design of the switching center equipment. It is also important that each of the number of degrees of precedence of a message be indicated by a specific precedence indicator. Both routing indicators and precedence indicators must appear in the routing line or address portion of the message.

Switching Center Specifications

The specifications covering the design of the switching center were highly exacting in their requirements. One of the most basic of the requirements was that the new switching center should take fullest advantage of existing teletypewriter equipment. This was necessary because the Army had a considerable amount of money already invested in teletypewriter apparatus which was usable. This requirement limited the new center to the type of apparatus to be used and fixed the network transmission speed at the capabilities of existing equipment, which was approximately 60 words per minute. Another basic requirement was that the new equipment should be capable of conforming to the then existing message format. It also was to be capable of installation at any switching center of the network without installation of similar equipment at other centers. These two requirements made it necessary for the new switching center to be capable of directly replacing any one of the switching centers in the network. The reason for this was that the first center was to be of an experimental nature, and the Signal Corps preferred not to permit the experimental switching center to compel alteration in the network procedure and message format in the entire world-wide network. From a practical standpoint, these requirements meant that the equipment must read the address portion of the message after it was perforated in a paper tape at the switching center and act upon it accordingly. One deviation from the standard message format which the Signal Corps permitted was the addition of a start-of-message indicator at the beginning of a message and an end-of-message indicator at the end of the message. Some commercial systems do not use a start-of-message indicator, but it was felt that in the Army system the start-of-message indicator was necessary to safeguard message sequence. Since a message might pass through several switching centers en route to its destination a further requirement was that none of the routing information could be absorbed in a switching center, thus insuring that succeeding switching centers would have complete routing information.

Because of the importance of military messages, the system had to be engineered for a maximum degree of reliability. Adequate provisions were to be incorporated into the system to insure that no message would ever become lost. The equipment was to be so well interlocked that equipment failures of any kind which might result in a lost or delayed message would be brought to the immediate attention of a supervisor. As a safeguard against lost messages in the early manual systems, each transmission of a message between two points was assigned a serial number which served as a basis for checking the receipt of all messages passing through the networks. A similar system was to be retained in the new system but was to incorporate automatic checking of the numbers.

General System Description

Fig. 1 shows a general view of the Fifth Army Headquarters switching center. The layout of equipment falls into two general classes. The first of these is the operational equipment; that is, those components which are normally used by operating and supervisory personnel in the regular performance of their duties. The second is the switching equipment which normally is unattended. There were several factors which influenced the decision to follow the plan of separation of operating and switching equipment. One of these was the type of equipment to be used. It was decided that electromagnetic switching equipment would be the most satisfactory, the most economical, and the easiest to use. A further consideration in the use of this type of equipment was that Signal Corps automatic-telephone central-office maintenance personnel would have the required background to learn quickly the circuitry of the new center.

The switching equipment can be basically divided into two classifications: the actual switches required to establish signal connections, and the relays required for tape reading, temporary data storage, and control purposes. The rotary switches are of two types. The first is a multilevel 25-point electromagnetically operated stepping switch, like that shown in Fig. 2. These switches are used primarily for the actual switching of signal circuits within the center, but are also used for control functions. The other type of switch used is the multilevel 10-point switch shown in Fig. 3. Switches of this type are used primarily for counting, since the 10-point continuous movement lends itself to cascading of switches to record a plurality of digits. They are used also in many places as a straight selector switch for control purposes.

The relays used are also of two types. The first of these, shown in Fig. 4, is the quick-acting multicontact type, using twin contacts for maximum dependability. The versatility of this relay was used to great advantage to obtain the desired operating characteristics of the circuits. Since much of the tape reading was accomplished on a character-by-character basis by these relays, they had to be fast in operation and release, and had to carry a large number of contact springs. A further consideration in the selection of this relay is its extremely long life. Since many relays in the system must operate once for each character received on any incoming line on a 24-hour basis, relays of rugged character were essential. Many relays in the

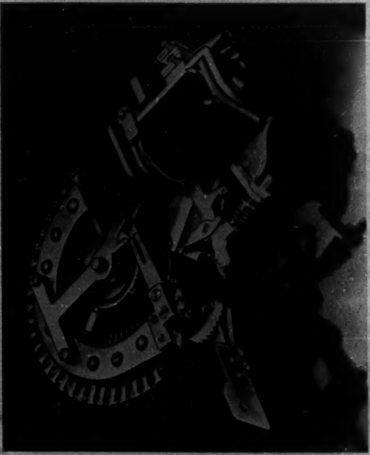


Fig. 2 (left). Twenty-five-point multilevel stepping switch

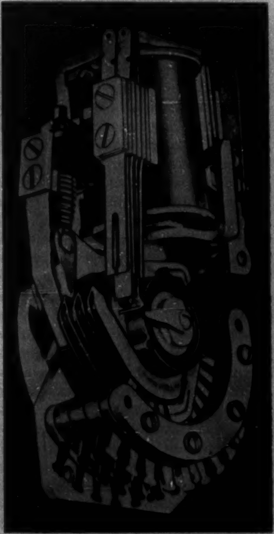


Fig. 3 (right). Ten-point multilevel stepping switch

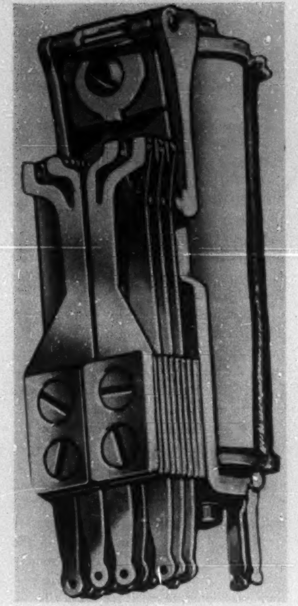


Fig. 4. General purpose multicontact relay

system have as many as 12 separate sets of contacts, and thus it was imperative to choose a relay capable of satisfactorily operating over long periods of time with such a large load.

The other type of relay used in the system is that shown in Fig. 5. It has roughly half the size and contact spring capacity as the other type, and two of these are mounted in a space equivalent to one of the others. This relay was chosen to conserve mounting space where necessary. It is ideally suited to use as a data-storage relay, since such relays generally require few springs and are used in large numbers.

A further consideration involved in the separation of the operational and the switching equipment was the decision to use common control equipment to the greatest possible extent. Here the principles which have been proved in automatic telephone exchanges were used to good advantage. Since it is basically uneconomical to use large amounts of equipment, required to establish a connection throughout the entire length of the message, the equipment used for this purpose was grouped together and made available to all components as required. Thus the total amount of switching equipment was materially reduced by the extensive use of common equipment. A further reason for the extensive use of common equipment was the fact that by its use, equipment quantity becomes a function of traffic rather than a function of the number of terminal facilities in the center.

Another factor which influenced the decision to separate operating and switching was the packaging requirements of the specification. The system was to be designed as a basic 25-line switching center capable of expansion to a maximum of 250 lines by the addition of similar components. The wiring of the banks of selector switches to a common multiple would be obviously uneconomical if the switches were located in different parts of the center, but when switches are packaged together, installation becomes simple, since most wiring has been done at the factory. The coupling together of additional switching cabinets to accommodate an expansion of the system is a relatively simple mattter. A switching unit of this type is shown in Fig. 6.

The operating components consist of:

1. Incoming line unit, Fig. 7;
2. Cross office unit, Fig. 8;
3. Automatic numbering and monitor unit;
4. Cross office intercept unit;
5. Incoming line intercept unit;
6. Intercept manual forwarding unit;
7. Multiple call storage unit;
8. Supervisor's console.

Also associated with the operating equipment but not directly coupled to it are teletypewriter devices for initiating new messages and keeping the necessary supervisory records.

Sequence of Operation

A general understanding of the operation of the system can be obtained by following the block diagram of Fig. 9, and by considering the flow of several messages through the center, which will illustrate the sequence of operations. The incoming lines pass through terminal and patching equipment (which is not shown) and terminate in typing reperforators in the incoming line unit. The typing reperforator reproduces the message on a paper tape both in typed form and in perforations in the tape. As the tape between the reperforator and the reader slackens, the tape reader senses each character in the tape and passes along an electrical indication of the characters to the associated control relay group. In this manner, the tape reader and the relay group co-operate in reading the tape. A relay sequence register in the relay group scans each character and records certain ones. When the sequence which is the start of message indicator arrives, the sequence register detects it and indicates its presence by the operation of a relay. The next bit of information appearing in the message format is the channel designation and number. It may be recalled that each incoming transmission on a given line is numbered serially. The channel number comparator has recorded the serial number of the last message received, thus being preset to record and check the serial number of the new incoming message. If the number in the comparator does not agree with the number on the incoming

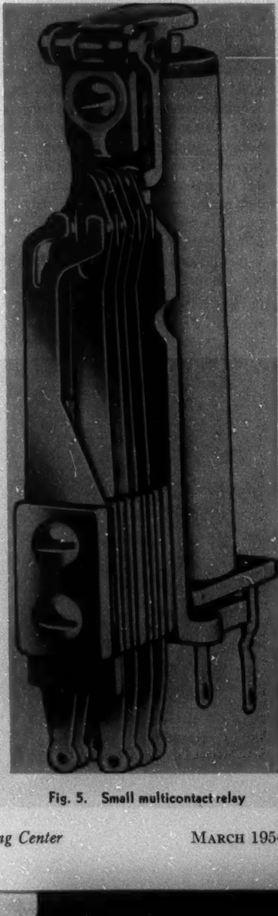


Fig. 5. Small multicontact relay

message, further processing is halted and an alarm condition is indicated. If, however, the numbers do agree, processing of the message is permitted to continue.

Checking of the channel number is a signal for the line circuit to call for the services of a director, Fig. 10. In response to that call, an idle director finds and associates itself with the calling line equipment. Following this, the reading of the tape by the reader is permitted to continue, and each character read is recorded on relays in the director. The director in scanning each character read by the reader, is hunting for one of several possible sequences of letters which will constitute a precedence indicator. Following receipt of this indicator other relay sequence registers are used to detect and register a routing indicator. These relay registers are unique in that they have the ability to detect a bona fide routing indicator from an indiscriminate group of characters, thus not relying on the relative position of the indicator in the message format.

When the routing indicator has been received, no further information is required from the line circuit. Receipt of the indicator also initiates a call by the director for the services of the routing translator. The routing translator is an all-relay device which receives from the director a coded 3-, 4-, or 5-character routing indicator, translates that information, and sends back to the director an indication of the outgoing line to be used for the message bearing that particular routing indicator. This translation process requires approximately 50 milliseconds, and since the translator is used for such a short time, it is common to all directors in any one switching center consisting of 100 lines or less.

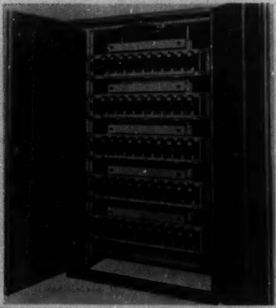


Fig. 6. Cross office selector switch unit

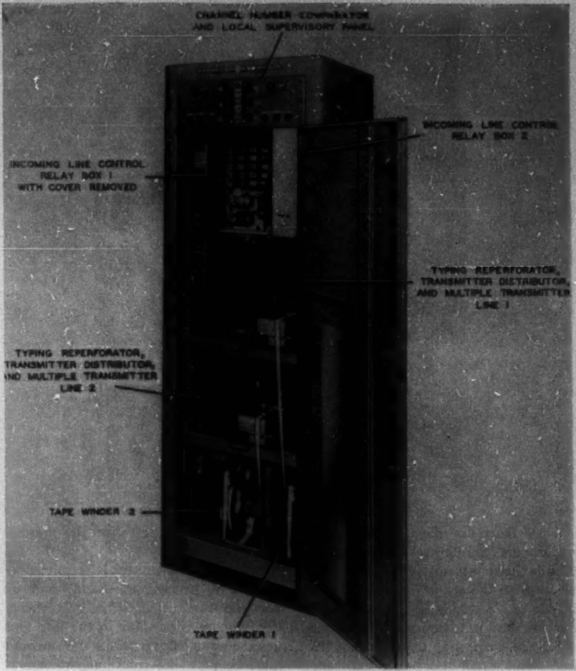

CHANNEL NUMBER COMPARATOR AND LOCAL SUPERVISORY PANEL
INCOMING LINE CONTROL RELAY BOX 2
INCOMING LINE CONTROL RELAY BOX 1 WITH COVER REMOVED
TYPING REPERFORATOR, TRANSMITTER DISTRIBUTOR, AND MULTIPLE TRANSMITTER LINE 1
TYPING REPERFORATOR, TRANSMITTER DISTRIBUTOR, AND MULTIPLE TRANSMITTER LINE 2
TAPE WINDER 2
TAPE WINDER 1

Fig. 7. Incoming line unit

Having received a line identification from the translator, the director must then select a cross office unit. These units are arranged in a common pool of equipment; that is, no unit is permanently connected to any incoming line unit or permanently connected to any outgoing line. A cross office unit is temporarily associated with an incoming line unit via the action of the cross office selector switch, which is associated with an incoming line unit, and it is temporarily connected to an outgoing line via the action of its own outgoing line selector switch. When a cross office unit has been associated with an outgoing line, it remains set to that line until all messages in it have been transmitted, following which its outgoing selector switch is reset and the unit is then returned to the pool and is again available for seizure for a new outgoing line. While a unit is set to an outgoing line it may receive many messages which are destined for that particular line providing the unit also has been set to the desired degree of precedence. Two messages of different degrees of precedence are never directed to the same cross office unit while the outgoing selector switch is set to a particular line.

During the process of selecting a cross office unit, the director may make several tests of the cross office pool. The first of these is to determine whether or not there is already at least one cross office unit preset to the desired outgoing line and the desired degree of precedence. If there is at least one such unit and it is not busy receiving a message from another incoming line unit, the director may cause the incoming line selector switch to be positioned to this unit. If the director does not find such a unit, one must be selected from the pool of units. The process of selecting a new unit involves setting the cross office selector switch to a unit which is available, and subsequently setting the precedence switch of the unit to the desired degree of precedence and also setting the outgoing line selector switch to the desired outgoing line. Following this the unit is ready to receive the transmission from the incoming line unit.

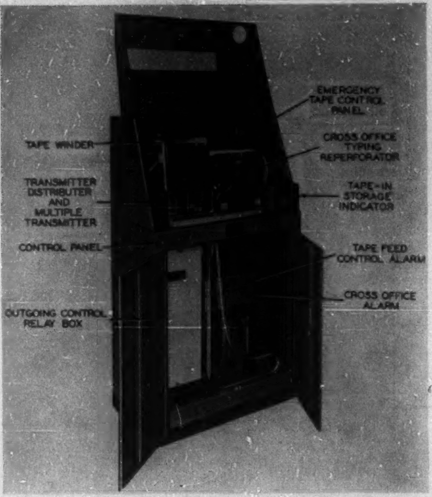

EMERGENCY TAPE CONTROL PANEL
CROSS OFFICE TYPING REPERFORATOR
TAPE WINDER
TRANSMITTER DISTRIBUTER AND MULTIPLE TRANSMITTER
TAPE-IN STORAGE INDICATOR
CONTROL PANEL
TAPE FEED CONTROL ALARM
CROSS OFFICE ALARM
OUTGOING CONTROL RELAY BOX

Fig. 8. Cross office unit

When the cross office unit has been properly set, the director has completed its functions. Its last function before release from the connection is the transmission of a signal to the incoming line unit to cause it to assume holding control of the connection to the cross office unit. This signal also causes the transmitter to commence transmission of the message to the cross office unit. Transmission to the cross office unit is at the rate of 75 words per minute, which is sufficiently faster than the 60-word-per-minute speed of the incoming reperforator to cause elimination of the tape which accumulated during switching. When the backlog of tape has been eliminated transmission continues at the rate of 75 words per minute, but on an intermittent basis.

When the start-of-message indicator has been transmitted to the cross office unit it is advanced immediately, in the case of seizure of a new cross office unit, to the tape reader, where the start-of-message indicator is read by the tape reader and detected by the associated relay group as in the case at the incoming line unit. After detection, the start-of-message indicator is advanced to a point immediately beyond the transmitter, at which time the reader and transmitter are stopped. At this time the associated relay group tests the condition of the outgoing line to determine whether or not the outgoing line is busy or idle. If found to be idle the line is immediately seized by the cross office unit, thus guarding the line against seizure by another unit.

Before the transmitter is permitted to proceed a transmission serial number must be transmitted to the line. Since this number is peculiar to the line and not the cross office unit the transmission must be controlled from line equipment. For this purpose, the automatic channel number transmitter is used to transmit first the start-of-message indicator, and then the channel designation and serial number. Immediately following transmission of the number, control of the line is returned to the cross office unit for transmission of the message. When the end of message indicator is sensed by the reader and detected by the relay group, the tape is advanced through the transmitter to a point where the end-of-message indicator has been advanced to, but not through, the transmitter. At this point the cross office unit releases from the connection and the automatic message-time transmitter transmits to the outgoing line the time of day followed by the end-of-message indicator.

In the example described above, the cross office unit found the outgoing line to be idle when tested, which resulted in immediate seizure of the line. Had the cross office unit found the line busy, a further test would have been made to determine the precedence assigned to the line by some other cross office unit. Having obtained an indication of this precedence, the cross office unit would have then made a comparison of the line precedence with its own precedence. If the precedence of the line at the time was low and the precedence of the cross office unit was sufficiently high (one of the three highest) the cross office unit would interrupt the outgoing message. It will be recalled that in a military network, certain messages must occasionally be given absolute priority to the line even at the cost of interrupting a message already in the process of transmission. In our example, when a cross office unit has made a decision to interrupt a message, a common piece of equipment known as a cancel transmission transmitter is temporarily connected to the line and transmits a cancellation sequence followed by an end-of-message indicator. Following this sequence a signal is sent to the cross office unit which had control of the line. This signal disconnects the cross office units control of the line, places it in a locked out condition and causes an alarm condition to be indicated. A supervisor must rethread the tape in the transmitter before permitting it to proceed. Procedure such as this occurs very infrequently, however. When the original cross office unit has been locked out, the line is immediately returned to an idle condition, thus permitting the new cross office unit containing the high precedence message to gain control of the line.

It is obvious that at the time that a line becomes idle there should be several cross office units waiting to seize it. In this situation all such cross office units first indicate precedence to the line. Electrical interlocks insure that only a cross office unit containing a message of the highest degree of precedence indicated will seize the line. Thus it can be seen that in a situation involving several cross office units all waiting for the same line message traffic can accumulate in the tape storage bins of these units to the capacity of the bins, each of which exceeds 20 average length messages. From a consideration of this situation it can also be seen that the number of cross office units set to a particular line is a function of the amount of traffic to that line. Since in a military network traffic volume to or from a particular station fluctuates widely, the pooled cross office units provide a completely flexible storage medium.

In a teletypewriter network, it is often necessary to repeat transmission. It is necessary, therefore, to keep a record of all transmission to each outgoing line. For this purpose a monitor unit is used. It consists of a typing reperforator which produces a record of transmission in punched tape form. These tapes are automatically taken up on reels and

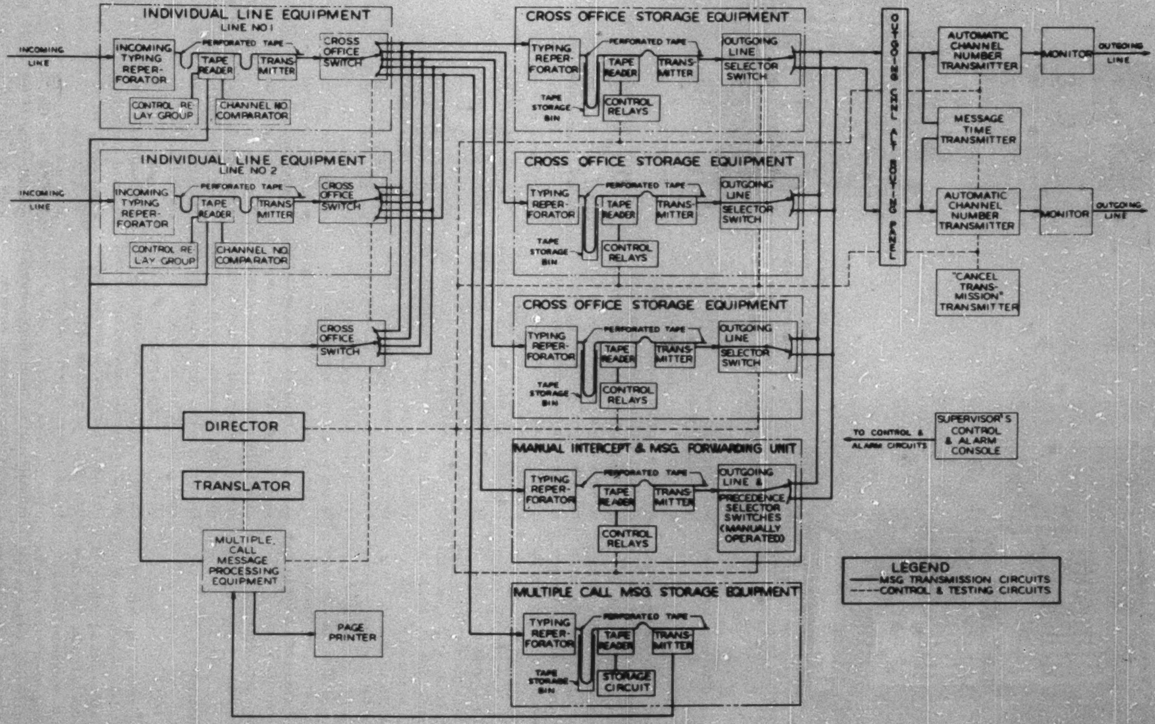

INDIVIDUAL LINE EQUIPMENT
LINE NO 1
INCOMING LINE
INCOMING TYPING REPER-FORATOR
PERFORATED TAPE
TAPE READER
TRANS-MITTER
CROSS OFFICE SWITCH
CONTROL RE-LAY GROUP
CHANNEL NO COMPARATOR
INDIVIDUAL LINE EQUIPMENT
LINE NO 2
INCOMING LINE
INCOMING TYPING REPER-FORATOR
PERFORATED TAPE
TAPE READER
TRANS-MITTER
CROSS OFFICE SWITCH
CONTROL RE-LAY GROUP
CHANNEL NO COMPARATOR
CROSS OFFICE SWITCH
DIRECTOR
TRANSLATOR
MULTIPLE CALL MESSAGE PROCESSING EQUIPMENT
PAGE PRINTER
CROSS OFFICE STORAGE EQUIPMENT
TYPING REPER-FORATOR
PERFORATED TAPE
TAPE READER
TRANS-MITTER
OUTGOING LINE SELECTOR SWITCH
TAPE STORAGE BIN
CONTROL RELAYS
CROSS OFFICE STORAGE EQUIPMENT
CROSS OFFICE STORAGE EQUIPMENT
MANUAL INTERCEPT & MSG. FORWARDING UNIT
OUTGOING LINE & PRECEDENCE SELECTOR SWITCHES (MANUALLY OPERATED)
MULTIPLE CALL MSG. STORAGE EQUIPMENT
STORAGE CIRCUIT
OUTGOING CHNL ALT ROUTING PANEL
AUTOMATIC CHANNEL NUMBER TRANSMITTER
MESSAGE TIME TRANSMITTER
AUTOMATIC CHANNEL NUMBER TRANSMITTER
"CANCEL TRANS-MISSION" TRANSMITTER
MONITOR
OUTGOING LINE
MONITOR
OUTGOING LINE
TO CONTROL & ALARM CIRCUITS
SUPERVISOR'S CONTROL & ALARM CONSOLE
LEGEND
MSG TRANSMISSION CIRCUITS
CONTROL & TESTING CIRCUITS

Fig. 9. System functional block diagram

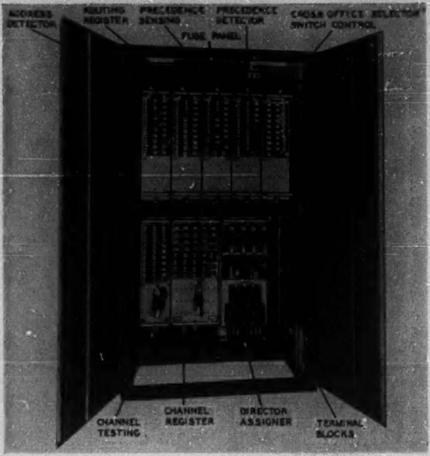


Fig. 10 (left). Director unit

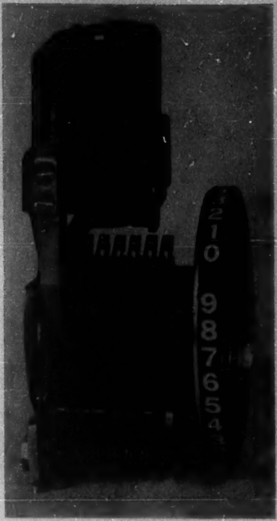


Fig. 11 (right). Channel number comparator switch showing use of the large indicator wheel

stored for a period of time for record purposes.

If a director fails to detect a bona fide routing indicator, receives a supervisory message for the local switching center, or for any reason cannot immediately direct a message to a cross office unit, the message will be directed to a cross office intercept position where the message is reproduced on a tape in such a manner that the tape can be easily torn between messages. This permits an operator to examine each message for appropriate action. If the message requires forwarding the operator has the facilities of the intercept manual forwarding unit by which the message can be reinserted into the network. Thus it can be seen that every message is immediately cleared from the incoming position to some form of cross office position. In a military network this is highly important, since the incoming line equipment must be left free to act upon a high precedence message. This is one of the fundamental departures from commercial techniques.

Equipment Description

The Incoming Line Unit

The incoming line unit, shown in Fig. 7, consists of a cabinet which houses all teletypewriter and switching equipment which is individual to two incoming lines. The teletypewriter equipment for each line consists of a typing reperforator equipped with a tape supply reel, a tape reader, a tape transmitter, and a tape take-up reel on which the spent tape is accumulated. The channel number comparators and other controls are located on the front panel. Inside are the two associated relay groups. In this case, the relay control groups are mounted with the operational equipment to facilitate maintenance and cabling of the center. The tape reader indicates to the relay group via five coding wires, which are grounded one or more at a time depending upon the character read. In this manner, the tape reader indicates a particular character to the relay group, where five sensing relays are used to decode the character. This application is an excellent example of the use of multicontact relays, since the decoding process requires indication of several characters, each of which requires two sets of contacts on each of the five relays.

With the receipt of the first characters of the start-of-message indicator, a set of two relays is operated, preparing the paths for operation of the second set of relays if the next character received is the actual second character of the start-of-message indicator. If the second character received is not the proper one, the first set of relays must immediately release so that it will be ready to operate on the receipt of the next first character received. Thus in this manner a sequence of several characters can be used and detected as a start-of-message indicator. Relays were chosen for this application instead of a stepping switch because a relay counting chain is capable of being instantly reset to the starting position if a character is received out of sequence.

One of the interesting features of the incoming line unit is the channel number comparator. It will be recalled that each transmission between two points carries a serial number and that the sequence of these numbers must be checked to guard against lost messages. This is accomplished by keeping in storage the serial number of the next message to be received, and comparing the stored number with the number on the incoming message. The stored number is registered on a set of three 10-point rotary switches of the type shown in Fig. 11. Each of these rotary switches is equipped with a large indicator wheel for displaying the contact number to which the switch is set. By mounting these three switches side by side behind a panel, with only one number of each switch showing through a hole on a panel, a 3-digit number can be indicated visually. This number is at all times the serial number of the next message to be received. The number is automatically advanced one unit as each message is processed. The banks of the rotary switch are used for automatic checking of the serial number. The visual indication is observed by the supervisor in case of an alarm condition resulting from failure to check a number. The actual comparison is accomplished by decoding the first character of the serial number and grounding one of ten leads. These ten leads are connected to bank contacts of the first switch. If the switch is set to the contact connected to the grounded lead a successful comparison is made and another character is read from the message tape. A fourth switch of the type shown in Fig. 3 is used to switch control to the second and third rotary switches for checking the succeeding numbers.

THE DIRECTOR

Broadly stated, the function of the director is that of directing the establishment of transmission paths through the center. This major objective is accomplished by proper control of the following functions:

1. Incoming line selection;
2. Precedence detection and storage;
3. Routing indicator detection and storage;
4. Control of translation and line identification storage;
5. Selection of a suitable cross office unit;
6. Setting of all necessary switches;
7. Release.

The director unit is shown in Fig. 10. It consists of two directors, one on each side of the cabinet. In the lower right-hand corner, the line selector switch base is visible. This base is common to the two directors, and consists primarily of two rotary switches per director mounted adjacent to each other and on the same base to facilitate wiring. The function of these switches is that of finding an incoming line position which has a message to be switched. Since these switches are 25-point switches, each set of two directors operates with 25 incoming line sets. Under normal traffic loads one of the two directors can adequately handle all traffic supplied by a group of 25 lines. This is a prime example of the savings realized by resorting to the use of common director equipment.

The precedence-sensing relay group is primarily a sensing means to supply the registers of the precedence detector. The output of the tape reader in the incoming line unit is connected to the five sensing relays here which decode characters received (which constitute precedence indicators) and pass them on to the detector registers of the predecence-detector relay base. In these relay registers, sets of 2-step relays are used as in the start of message detector in the incoming line unit to pick a particular sequence out of a group of characters. Here, as before, relays were used to provide an immediate reset in case a character is received out of sequence.

The address-detector relay group is primarily a set of sensing relays similar to those of the precedence-sensing relay group, but in this case the contacts of the relays are wired to decode only certain letters of the alphabet which are used as first or second characters of the routing indicators. The output of this unit is fed to the routing-register relay base, which is a means of storing temporarily a 3-, 4-, or 5-letter routing indicator in code form. The operation of the routing register relay base is of particular interest. It consists of five sets of five storage relays each and five control relays arranged in a counting chain for control of the registers. In the routing-indicator assignment plan, the first and second characters each can be one of only a selected few letters, and each of the remaining characters can be any one of the letters of the alphabet. Thus a receipt of two characters which are bona fide letters for their respective positions will establish the basis for a routing indicator sequence. These 3, 4, or 5 characters are then stored in the storage registers in sequence.

One of the interesting features of this piece of equipment is its ability to pick a routing indicator out of a group of indiscriminate characters. The receipt of a usable first character causes that character to be stored in the first relay register. The receipt of a usable second character is stored in the second set of registers, thus preparing the third and subsequent registers to receive the remaining characters. If, however, a usable first character is received which is not followed immediately by a usable second character, the first and second registers must be immediately released between characters to be capable of receiving a new first character should one appear. A further aspect of this situation appears when a usable first character has been received which is not usable as a second character but happens to be usable in the first position as the start of a new sequence. In this case the first register must be released and the second character must be stored as a first character. The use of multicontact relays here facilitates the rapid switching required to store the coded characters in the proper registers.

THE ROUTING TRANSLATOR UNIT

One of the most basic functions of the switching equipment is to determine which outgoing line to select for a particular message. This function is accomplished in approximately 50 milliseconds by the routing translator, which is common to all directors in any one switching center consisting of one hundred lines or less. The translation processes are accomplished by means of multicontact relays of the type shown in Fig. 4, and by means of patching cords with pin-jack connections. In a commercial system, network changes are relatively infrequent. In a military network, however, overnight changes can and do occur very frequently. It is also possible that military routing indicator assignments may change completely overnight. These factors have dictated the use of pin-jack patching cords in the translation process to permit rapid changes of the switching plan. The translator is also supplied in duplicate, which permits operation on one switching plan with immediate transfer to another plan when required.

The basic translator circuitry has been diagrammed in Fig. 12. Information is supplied from the director via the five sets of five leads each to operate the five sets of decoding relays. Each of these sets of decoding relays decodes an encoded letter and indicates the letter by operation of one of its associated translation relays. It will be recalled that the first three letters of a routing indicator identify a particular relay center in the network. The operation of one of several possible first position translation relays extends a ground to contacts, in multiple, of all second position relays. The operation of one of the second position relays likewise extends one grounded lead to contacts of all third position relays. The operation of one of the third position relays extends the ground to a pin jack which identifies a particular relay center. Likewise all possible combinations of the first three characters are connected to similar pin jacks. The arrangement of switching centers and tributary stations is such that no switching center ever need transmit a message directly to a tributary station of another switching center. Thus the receipt of a routing indicator which has the first three letters identifying a center other than the local switching center does not require action on the fourth or fifth character. Thus the pin jack in the 3-letter field which represents the local switching center is strapped to the "local" pin jack, which starts another translation ground through the fourth and fifth character translation relays, ultimately to reappear at the fourth or fifth character pin-jack fields. Single-conductor patching cords are used to connect the assigned routing indicators of the network to the desired outgoing line.

The Multiple Call System

Before we consider the equipment required to process multiple-address messages, we should have some understanding of the multiple-address procedure as

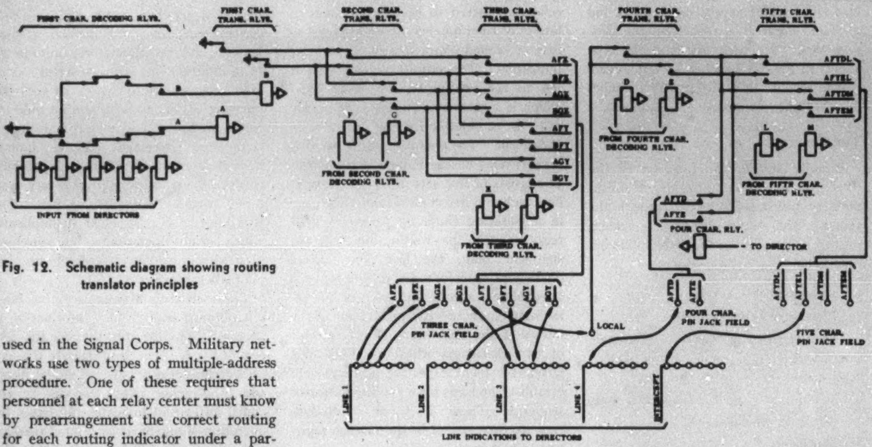

FIRST CHAR. DECODING RLYS.
FIRST CHAR. TRANS. RLYS.
SECOND CHAR. TRANS. RLYS.
THIRD CHAR. TRANS. RLYS.
FOURTH CHAR. TRANS. RLYS.
FIFTH CHAR. TRANS. RLYS.
A
B
INPUT FROM DIRECTORS
F
G
FROM SECOND CHAR. DECODING RLYS.
AFX
BFX
AGX
BGX
AFY
BFY
AGY
BGY
X
Y
FROM THIRD CHAR. DECODING RLYS.
D
E
FROM FOURTH CHAR. DECODING RLYS.
AFYDL
AFYEL
AFYDM
AFYEM
L
M
FROM FIFTH CHAR. DECODING RLYS.
AFYD
AFYE
FOUR CHAR. RLY.
TO DIRECTOR
THREE CHAR. PIN JACK FIELD
LOCAL
FOUR CHAR. PIN JACK FIELD
FIVE CHAR. PIN JACK FIELD
LINE 1
LINE 2
LINE 3
LINE 4
INTERCEPT
LINE INDICATIONS TO DIRECTORS

Fig. 12. Schematic diagram showing routing translator principles

used in the Signal Corps. Military networks use two types of multiple-address procedure. One of these requires that personnel at each relay center must know by prearrangement the correct routing for each routing indicator under a particular set of circumstances and must route all messages accordingly. This type of routing is used to best advantage when the network is constructed along command lines or is so arranged that the bulk of message traffic is from or to a centralized switching center with little or no traffic passing between adjacent switching centers. The other type of multiple-message processing requires that personnel at the originating station assign a definite responsibility to a particular relay center. This method is used by the Signal Corps primarily because it avoids the confusion of considering message origin in networks which make extensive use of interconnecting trunks between many switching centers of the network. As an example of this type of message we can assume a message originating at station *AA* which is a tributary station of switching center *A*. An operator at station *AA* addresses a message to station *CA* and *CB*, and instructs relay center *C* that it is to be responsible for multiple transmission. The message may travel as single transmissions through relay centers *A* and *B* because only relay center *C* has been designated as the one to make a multiple transmission. This method requires knowledge of the network by the originating operator and by switching center operators, but it has the advantage of fixing definite responsibility for transmission of the message. Quite frequently a switching center will receive a multiple-address message, and in segregating the routing indicators for retransmission, finds that two or more routing indicators can be reached via another switching center. In this case a pilot is attached, preceding the heading of the message, assigning further multiple-address responsibility for the routing indicators concerned. In all cases, a pilot preceding a message is followed immediately by another pilot or by the original message heading.

Sequence of Operations

When a director detects the special sequence indicating local responsibility for a multiple-address message, the message is routed immediately to a multiple-call storage unit. This unit is similar in appearance to the cross office unit, and serves to collect multiple-address messages prior to processing. As in the case of the cross office unit the tape is read immediately by a tape reader, which transfers coded characters to the associated relay control group. When the start of message indicator is sensed, the tape reader is stopped until the processing equipment is available. The processing equipment is common to several storage units, and is assigned to each for the duration of one message.

When the processing equipment is connected to the storage unit, the tape reader is permitted to continue, and information passes from the tape reader to the processing equipment. The precedence of the message is recorded in the processing equipment, following which the first routing indicator is detected and stored. The receipt of the first routing indicator causes the tape reader to stop and start a group of operations. The first of these operations is a translation, and for this purpose the routing translator is used. The second operation is the placing in storage of the associated routing indicator. Perforated paper tape was chosen for these storage mediums because it permits use of up to 100 routing indicators in any one message. The remaining operation is the seizure of the cross office unit which will be held for transmission at a later time. A cross office unit is selected for this purpose in exactly the same manner as in the case of single-address message.

Following the completion of these operations, the second and all subsequent routing indicators are read and processed similarly. Thus a new cross office unit is seized each time a different outgoing line is indicated. Part of the processing involves keeping in relay storage the three letters identifying the relay center which is to be responsible for multiple processing of any of the outgoing transmissions. Thus when a particular outgoing line is indicated by the routing translator for a second or subsequent time, a comparison of the new routing indicator and the one previously in storage must be made. As a result of this comparison, a new routing indicator might be placed in storage in place of the one previously held. Thus for each indicated outgoing line to be used, the 3-letter switching center routing indicator is held in storage for use at a later time.

After the last routing indicator has been processed, transmission to the cross office units is started. The first item of information is transmitted simultaneously to all seized cross office units. It consists of the start-of-message indicator and the precedence indicator. Following

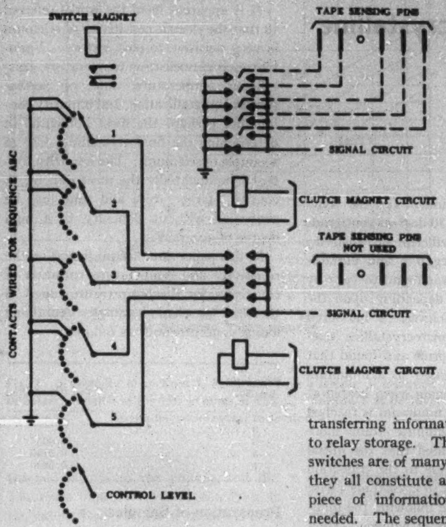


Fig. 13. Schematic diagram showing the markable transmitter principles

this, individual pilots are transmitted to each cross office unit as required. These pilots consist of individual handling instructions assigning specific transmission responsibility where required. For this purpose, the 3-letter switching center designations mentioned above become part of the respective pilots. Following this the routing indicators, which were previously placed in tape storage, are withdrawn from storage and transmitted to the cross office units, which are indicated by the line numbers, which were also placed in storage. In this manner a specific pilot, including applicable routing indicators, is transmitted to each cross office unit. Following completion of all of these pilot messages, the original address portion and the body of the message are transmitted from the multiple call storage unit in multiple to all seized cross office units. Following completion of the message, the processing equipment and the cross office units are released.

The multiple-address processing equipment is largely a collection of relays and rotary switches which are connected together to control a logical sequence of events. This equipment can be grouped into the following general classifications:

1. Sensing relays;
2. Storage relays and switches;
3. Transmission control switches;
4. Sequence control relays and switches.

The sensing relays are a means of transferring information from tape form to relay storage. The storage relays and switches are of many different forms, but they all constitute a means of putting a piece of information away until it is needed. The sequence control relays are a means of controlling all of the functions of sensing, storage, and transmission. The transmission control switches are a means of encoding information taken from storage to permit its transmission as required.

The sensing relays have been previously described. The encoding equipment, however, consisting of the transmission control switches, is equally important to the functioning of the multiple call system, because one of the functions of the system is the creation of specific pilot messages which must be transmitted as a sequence of teletypewriter characters. The basic principles of encoding are illustrated by the schematic diagrams of Fig. 13. The upper diagram shows a schematic representation of a typical teletypewriter transmitter. It consists of a set of six transmitting contacts which are wired as shown. When nothing is being transmitted, a ground is maintained on the signal lead via the normally closed contacts. As a mechanical cycle starts, these contacts open for a definite period, following which the remaining five contacts close or remain open in timed sequence, depending upon the character to be transmitted. The action of the contacts is controlled by the positions of the sensing pins, which in turn are controlled by the presence or absence of holes in the tape.

The multiple-call system requires in many instances that information be withdrawn from relay or rotary switch storage and transmitted directly to the signal circuit. To accomplish this function, a standard transmitter was modified by splitting apart the common electrical connection on one side, and wiring it to a control switch as shown in the lower diagram. By operating the transmitter without tape all sensing pins are permitted to operate, thus causing all contacts to close in sequence. External control of the grounds to the respective contacts has the same effect as closing or opening the contacts through operation of the sensing pins in a normal transmitter. In operation, a pulse is sent to the rotary switch, which steps from its normal position to its first position. Immediately following this a pulse is applied to the clutch magnet of the transmitter, thus causing transmission of the character which is represented by the ground pattern on the first five levels of the switch. Subsequently, succeeding pulses are applied to the switch and the transmitter, thus causing a multiletter sequence to be transmitted. It should be noted that, while the control switch as shown has its bank contacts grounded directly to transmit a fixed sequence, these contacts are often connected to other supplementary switches, thus permitting transmission of a variable sequence. This plan of controlling a teletypewriter transmitter is the heart of the multiple-call processing operation.

Conclusion

This experimental installation at Fifth Army Headquarters has been regarded as a first step towards the goal of a completely automatic network. Several larger switching centers are being planned which will incorporate many new features and techniques, as well as a completely new line of teletypewriter equipment which is being adapted especially for use with automatic-switching techniques. Much effort is being exerted to streamline network procedure and message format. It is expected that these efforts and the more extensive use of automatic switching equipment such as this will result in smoother and more efficient operation in meeting the ever-increasing demand for more and better communications.

Reference

1. A Full Automatic Private-Line Teletypewriter Switching System, W. M. Bacon, G. A. Locke. *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 473–80.

◆

No Discussion

Electrical Properties of Microcrystalline Selenium

GILBERT HALVERSON
NONMEMBER AIEE

THE increasing number of electronic applications for semiconducting materials which have been developed in recent years, coupled with the discovery of the transistor action of these materials, has been largely responsible for the vast amount of research data which have been published on semiconductors. Most of the work has been done on germanium and silicon and as a result these data have led to a better understanding of the mechanism of electric conduction in semiconductors as well as an explanation of their rectification properties. Data such as these are lacking for selenium. The development of the selenium rectifier industry, however, during the past decade has stimulated an interest in its semiconducting properties. A limited amount of research work has been done on both single crystal and microcrystalline selenium, thinking that perhaps such data would aid in understanding the mechanism of conduction and also aid in overcoming some of the difficulties encountered in the manufacture of rectifiers. Several researchers have studied the electrical properties of single crystals of selenium with the thought that such data would serve to explain the electrical properties of microcrystalline selenium. These data have not materially clarified the situation. In view of these facts, and also since selenium exists in the microcrystalline state in selenium rectifiers, it would seem that measurements on the electrical properties of microcrystalline material would be the most practical contribution to existing data.

Measurements on the resistivity of microcrystalline selenium have recently been made.[1-4] Blackburn[1] measured the resistivity and the temperature coefficient of resistivity of selenium containing various percentages of bromine and reports a minimum in the resistivity for 0.007-per-cent bromine. In the temperature range of −30 degrees centigrade (C) to 100 C the coefficient of resistivity was found to be negative and variable and the resistivity was found to vary by factors of 3 to 12, depending upon the bromine content. Schweikert[2] measured the resistivity of microcrystalline selenium containing chlorine and found that the resistivity decreased to a minimum and then increased again when heated at 210 C, and that the minimum is reached earlier when the selenium contains more chlorine. He concluded that the initial rapid decrease was due to the transition from the amorphous to crystalline phase and the slow increase following the minimum was due to loss of halogen. Henkels[3] has reported extensive data on resistivity as a function of heat treatment and halogen content. He found that the resistivity depends only to a slight extent on the nucleation temperature and to a much greater extent on the crystallizing temperature. Plessner[4] found that the resistivity of samples crystallized at 110 C followed by 40 minutes at 218 C was lower by a factor of 10 than a similar sample heated at 175 C instead of 110 C. He attributes the difference to variations in crystal size, the higher temperature producing the largest crystals.

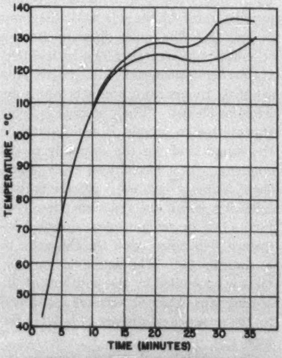


Fig. 1. Heating curves illustrating the temperature rise of selenium during the hot pressing operation

It is apparent from the works referred to that the electric resisitivity of selenium is very sensitive to past history. Variables such as nucleation temperature, crystallizing temperature, time of nucleation and crystallization, and impurity content are perhaps the most influential in determining the final crystalline state of a sample of selenium. The impurity content is undoubtedly the most difficult to control; temperature and time may be controlled without difficulty to a high degree of accuracy.

In this paper data are presented on the resistivity and temperature variation of resistivity for the temperature range 25 to 150 C on selenium samples containing iodine concentrations as follows:

Selenium Lot No.	Per Cent Iodine
0	0
5	0.0055
3	0.0190
1	0.0860

Preparation of Samples

The selenium used in preparing the samples was obtained from the American Smelting and Refining Company. No attempt was made to purify the selenium to a higher degree of purity. In preparing each sample for resistivity measurements 10 grams of selenium was heated for 30 minutes at 475 C in a clean porcelain crucible; it was then quenched by pouring into a special chromium plated steel die which had been cooled to 12 C. The die was then placed between the electrically heated platens of a hydraulic press and a pressure of 5 tons applied; the platens were thermostatically controlled at a temperature of 135 C. Each sample was left at this pressure and temperature for a total of 35 minutes after which the pressure was released, the die removed from the press, dismantled, and the sample removed. This procedure resulted in samples 5 inches by 0.5 inch by 0.040 inch which were uniform in cross section, and free from voids and cracks on the surface and throughout the interior as nearly as could be ascertained.

The temperature cycling of the selenium during this pressing operation is shown in Fig. 1. The temperature in each case was obtained by inserting a thermocouple in the side of the die such that it was located within 1/32 inch of the selenium. Approximately 20 minutes were required for the selenium to reach the temperature of the platens of the press. Fluctuations in temperature from 20 minutes to 35 minutes were due to

---

Paper 54-163, recommended by the AIEE Metallic Rectifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 21, 1953; made available for printing December 4, 1953.

GILBERT HALVERSON is with the Fansteel Metallurgical Corporation, North Chicago, Ill.

The author wishes to thank Mrs. Mary Conolly for preparation of the photomicrographs.

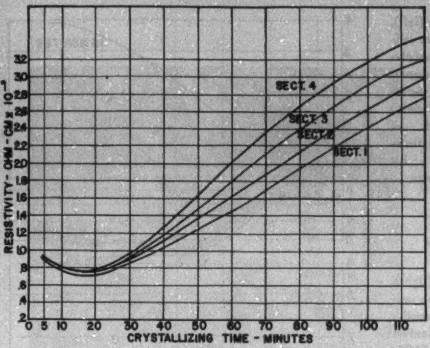


Fig. 2. Resistivity of sections 1, 2, 3, and 4 for a sample of undoped selenium crystallized for 120 minutes at 212 C. Section 1 at positive potential with respect to section 4

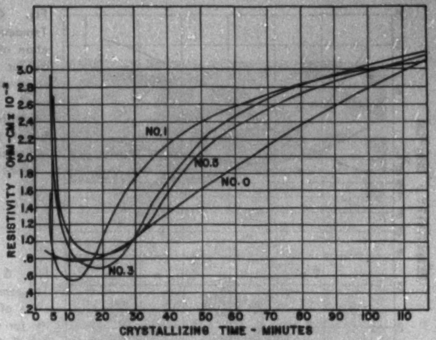


Fig. 3. Resistivity at 212 C as a function of crystallizing time for different iodine concentrations. Curve no. 0, 0-per-cent iodine; curve no. 5, 0.0055-per-cent iodine; curve no. 3, 0.0190-per-cent iodine; curve no. 1, 0.086-per-cent iodine

thermal inertia of the platens and die. The cycle, however, was reproducible to the extent that variations of ±5 C were seldom exceeded.

Resistivity data were taken on each sample during the crystallizing process. This was accomplished by inserting each sample into a special fixture containing two current probes and five potential probes spaced at 2-centimeter intervals along the 5-inch sample. With five potential probes it was possible to measure the potential drop, and hence resistivity, on four separate sections of each sample. In this manner any inhomogeneities in the structure would be detected. Silver was evaporated onto the selenium at each probe-selenium contact point; this lowered the contact resistance and substantially increased the sensitivity with which potential measurements could be made. The fixture and sample was then placed in a constant temperature oven at 212 C. A constant potential of 25 volts d-c was immediately applied across the sample. Readings of current and potential across each of the four sections were taken at intervals of time during this crystallizing process, the initial reading having been taken 5 minutes after placing the sample in the oven. Five samples were made from each of the lots of selenium listed here and these were crystallized at 212 C for total times of 15, 35, 60, 90, and 120 minutes respectively. That is, one sample was heated for 15 minutes and removed from the oven, the second was heated for 35 minutes, etc. Potential measurements were taken across each of the four sections of each sample with a Leeds & Northrup type-*K* potentiometer; current readings were take with a model-*430* Weston milliameter.

The temperature variation of resistivity was measured for each sample over the range 25° to 150 C. This was done by placing the sample in the oven at 25 C and raising the temperature step-wise, allowing 1/2 hour for thermal equilibrium at each temperature setting. In this manner each sample received the same thermal treatment so that changes in microcrystalline structure due to continued crystallization during this heat treatment would be identical for all samples. No attempt was made to carry the samples through several temperature cycles to study the hysteresis effect observed by Henkels.[6]

Experimental Results

Resistivity data in ohm-centimeters at 212 C are shown in Figs. 2 and 3. The data for each of the four sections of each sample were plotted as a function of crystallizing time as shown in Fig. 2. These data are for sample no. 5 of selenium lot no. 0 (undoped selenium) which was crystallized for 120 minutes. An average curve was drawn for these data. A similar procedure was repeated for each of the other four samples from this lot of selenium. The resulting curves representing the average resistivity values for the four sections of each sample were combined into a single curve which is shown for each of the four lots of selenium in Fig. 3.

The temperature variation of resistivity in the range from 25 to 150 C are shown in Figs. 4, 5, 6, and 7 for selenium lots no. 0, 5, 3, and 1 respectively. Values of resistivity in ohm-centimeters are plotted as a function of reciprocal of absolute temperature $1/T$. These data show the variation in the temperature coefficient of resistivity with time of crystallization at 212 C.

Photomicrographs were prepared from representative sections of each of the samples for purposes of illustrating the effect of crystallizing times on the microcrystalline structure. These are shown in Figs. 8, 9, 10, and 11 for selenium lots no. 0, 5, 3, and 1 respectively.

Discussion

The data in Fig. 3 showing the resistivity at 212 C as a function of crystallizing time agree qualitatively with the results reported by Schweikert[2] for chlorine doped selenium. The resistivity decreases, passes through a minimum, and then gradually increases again. Schweikert found minimum values ranging from 300 to 1,000 ohm-centimeters in about 100 minutes at 210 C. Our data indicate minimum values from 500 to 850 ohm-centimeters in 10 to 20 minutes as 212 C indicating that the crystallizing process is much faster with iodine as an impurity than it is for chlorine.

The rapid decrease in resistivity during the first few minutes at 212 C is attributed to the transformation from the amorphous to crystalline phase, the transformation being more rapid for the higher iodine concentrations. Continued heating with the observed increase in resistivity is attributed to increase in grain size. If

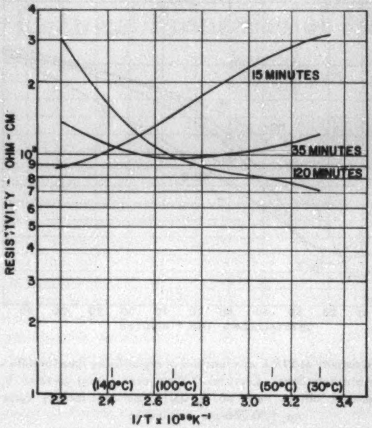


Fig. 4 (left). Temperature variation of resistivity for undoped selenium (0-percent iodine) crystallized at 212 C for times indicated by each curve

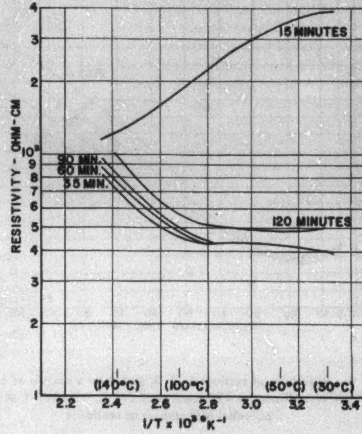


Fig. 6 (right). Temperature variation of resistivity for selenium no. 3 (0.0190-per-cent iodine) crystallized at 212 C for times indicated by each curve

crystallization is a purification process the impurity concentration at the boundaries of the grains increases as the grains continue to grow. Therefore, if conduction takes place along grain boundaries as was postulated by Henkels,[6] we would expect to find a decrease in resistivity. However, due to the increased size of the grains the number of conduction paths is considerably reduced, tending to increase the resistivity. This increase must more than compensate for the decrease due to increased impurity concentration in the grain boundaries with a net result of an increase in resistivity.

The data on temperature coefficient of resistivity are not sufficient by themselves to allow any theoretical interpretation; other data such as mobility and concentration of current carriers are necessary in order to interpret these data in the manner that Pearson and Bardeen,[7] for example, have done for silicon and germanium. It is interesting to note, however, that both heat treatment and impurity content make an appreciable change in the coefficient.

The progress of crystallization and grain growth is well illustrated by the series of photomicrographs in Figs. 8, 9, 10, and 11. From these photographs it appears that the presence of iodine retards the formation of the dendritic structure characteristic of crystallization from a melt and which is observed, for example, in plate 3849 of Fig. 8. The dendritic structure observed in this plate tends to decrease with increasing iodine concentration as may readily be seen from plates 3841, 3836, and 3831. Increasing amounts of iodine speed up the grain growth process during heating times from 15 minutes to 60 minutes; after 60 minutes there appears to be very little significant difference in the samples. If resistivity is dependent on grain size then the data in Figs. 4, 5, 6, and 7 further substantiate this observation. There is a marked drop in the room temperature resistivity when increasing the crystallizing time from 15 to 35 minutes; this change is less when increasing the time to 60 minutes, and after 60 minutes there is very little difference. Large percentages of iodine impurity, however, change this picture somewhat as may be seen in Fig. 7. The resistivity is a minimum for a crystallizing time of 35 minutes; longer periods of

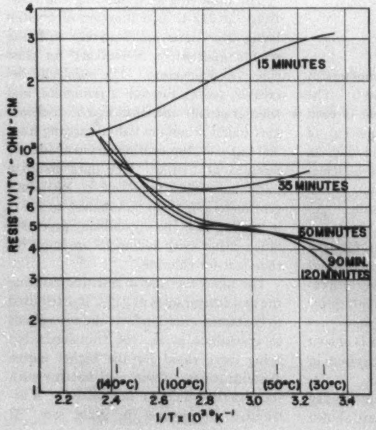


Fig. 5 (left). Temperature variation of resistivity for selenium no. 5 (0.0055-per-cent iodine) crystallized at 212 C for times indicated by each curve

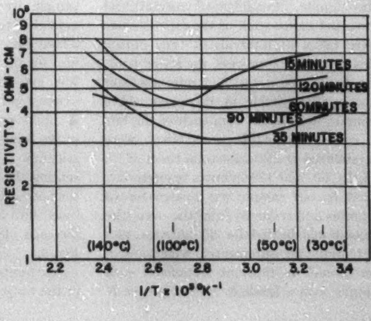
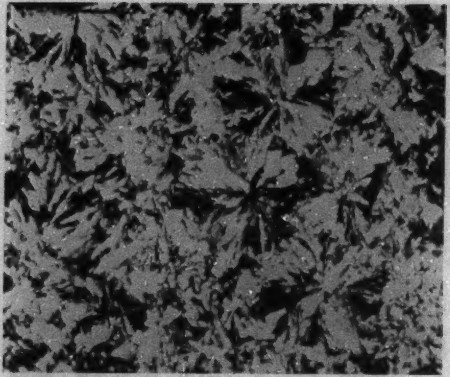


Fig. 7 (right). Temperature variation of resistivity for selenium no. 1 (0.086-per-cent iodine) crystallized at 212 C for times indicated by each curve

Fig. 8. Photomicrographs of selenium no. 0 (0-per-cent iodine). Magnification 600 times. Crystallized at 212 C as follows:



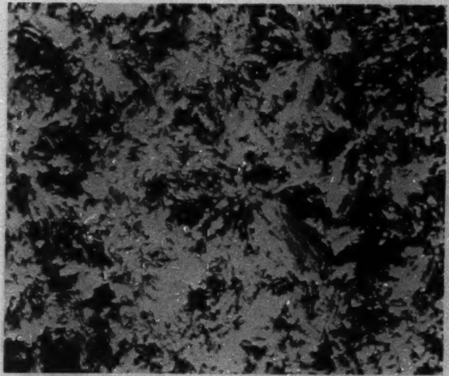
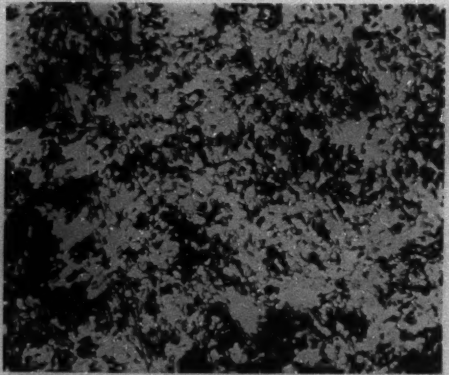


Plate 3849 (left), 15 minutes
Plate 3850 (right), 35 minutes



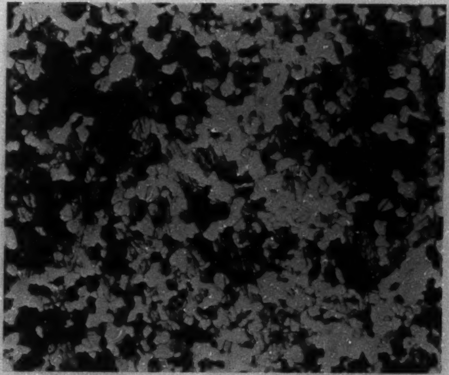
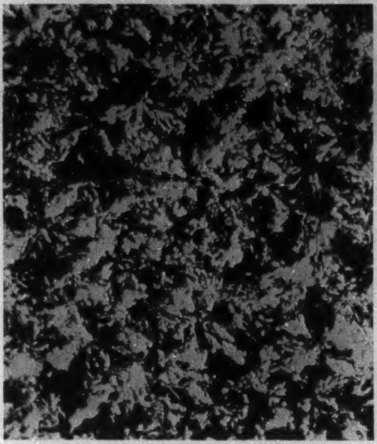


Plate 3851 (left), 60 minutes
Plate 3854 (right), 120 minutes

Fig. 9. Photomicrographs of selenium no. 5 (0.0055-per-cent iodine). Magnification 600 times. Crystallized at 212 C as follows:



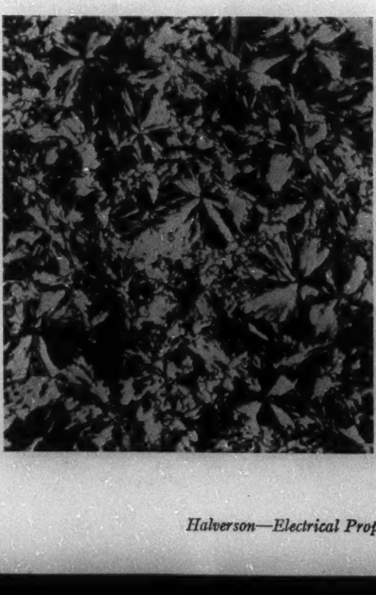
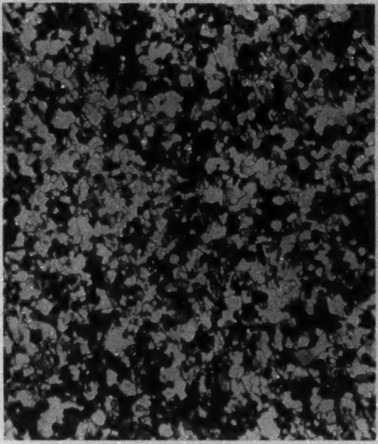


Plate 3841 (left), 15 minutes
Plate 3842 (right), 35 minutes



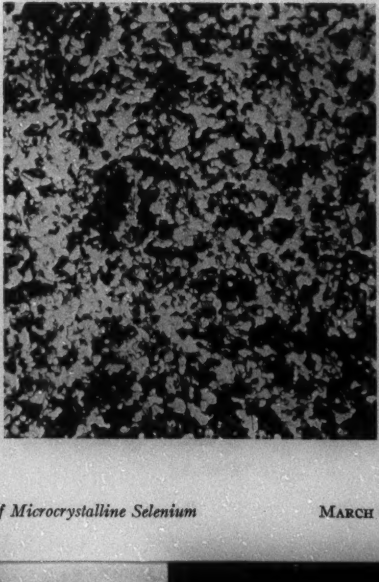


Plate 3843 (left), 60 minutes
Plate 3845 (right), 120 minutes

Fig. 10. Photomicrographs of selenium no. 3 (0.0190-per-cent iodine). Magnification 600 times. Crystallized at 212 C as follows:

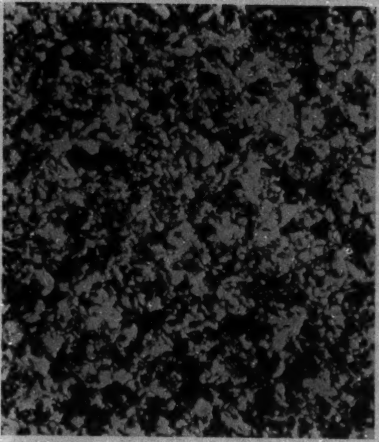


Plate 3836 (left), 15 minutes
Plate 3837 (right), 35 minutes

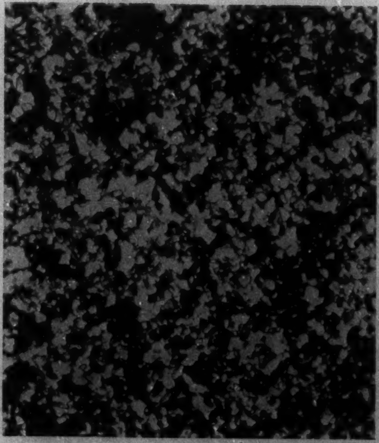
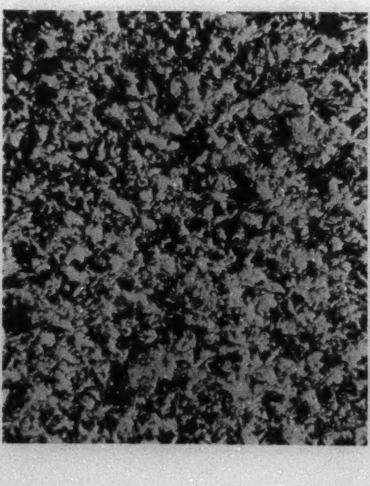


Plate 3838 (left), 60 minutes
Plate 3840 (right), 120 minutes

Fig. 11. Photomicrographs of selenium no. 1 (0.086-per-cent iodine). Magnification 600 times. Crystallized at 212 C as follows:



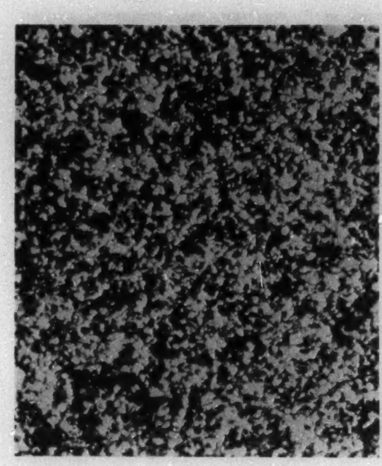
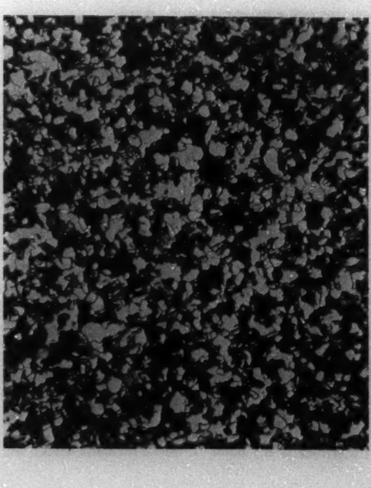


Plate 3831 (left), 15 minutes
Plate 3832 (right), 35 minutes



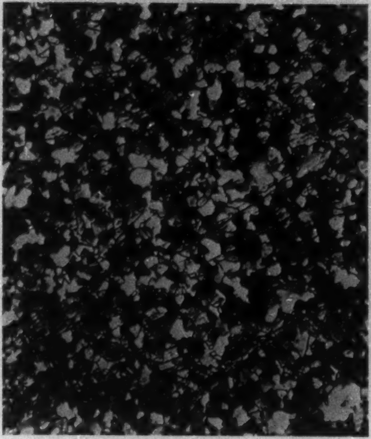


Plate 3833 (left), 60 minutes
Plate 3835 (right), 120 minutes

crystallizing time increase the resistivity. There is no apparent dependence of grain size on impurity content when the samples are crystallized for 120 minutes. This is shown by the similarity of plates 3854, 3845, 3840, and 3835 of Figs. 8, 9, 10, and 11, respectively.

There is a definite correlation between the resistivity at 212 C and the state of crystallization. The most rapid change in resistivity after the initial rapid decrease occurs within the interval of 15 to 40 minutes, as shown in Fig. 3. The photomicrographs in Figs. 8, 9, 10, and 11 show that the largest change in structure occurs for crystallizing times from 15 to 35 minutes. Fig. 3 shows that the resistivity values tend to converge to a common value at 120 minutes crystallizing time which substantiates the observation made here regarding similarity of structure after 120 minutes at 212 C.

Conclusions

1. The crystalline state of commercial grade selenium is markedly dependent upon the thermal treatment it receives.

2. Iodine impurity additions accelerate the crystallization process when the selenium is heated at 212 C.

3. The final state of crystallization, in particular the grain size, is independent of iodine content after crystallizing times of 120 minutes at 212 C.

4. Temperature coefficients of resistivity are both positive and negative and variable over the temperature range from 25 to 150 C.

References

1. **W. E. Blackburn.** *Thesis*, University of Pittsburgh, Pittsburgh, Pa., 1939.

2. **Hans Schweikert.** *Zeitschrift fuer Physik*, Berlin, Germany, vol. 128, 1950, pp. 47–55.

3. **Herbert W. Henkels.** *Technical Report No. 25*, United States Bureau of Ships, Washington, D. C., Contract NObsr-42487, March 13, 1951.

4. **K. W. Plessner.** *Proceedings*, Physical Society, London, England, vol. 64, pt. B, 1951, pp. 681–90.

5. **Herbert W. Henkels.** *Technical Report No. 14*, United States Bureau of Ships, Washington, D. C., Contract NObsr-42487, December 1, 1951.

6. **Herbert W. Henkels.** *Technical Report No. 19*, United States Bureau of Ships, Washington, D. C., Contract NObsr-42487, February 6, 1950.

7. **G. L. Pearson, J. Bardeen.** *Physical Revue*, New York, N. Y., vol. 75, 1949, p. 865.

No Discussion

Instantaneous Electrical Characteristics of Selenium Rectifiers

G. F. PITTMAN, JR.
ASSOCIATE MEMBER AIEE

IN POWER applications of selenium rectifiers the reverse characteristics of the selenium cell are of interest only in so far as they affect the heat generated within the cell and thus the efficiency and the current rating of the cell as dictated by its heat-dissipation capabilities. In more critical applications, however, the detailed reverse characteristics of the cell must be considered; this is particularly true in the case of selenium cells used as self-saturating or feedback rectifiers of magnetic amplifiers. In such applications, not only a more complete knowledge of the reverse resistive properties of the cell is required but also information regarding its capacitive properties.

Through an experimental study of the instantaneous reverse voltage-current operating loci of selenium cells at power frequencies, much information as to the nature and magnitudes of these reverse characteristics may be obtained. This information, in turn, provides an insight into the effects of these characteristics upon the operation of selenium rectifying elements and, also, upon methods of testing cells to be used in critical applications. This paper presents some results of such a study; relatively long-term effects such as aging, temperature drifts, forming, etc., are not considered, but, rather, attention is focused upon phenomena taking place within a given cycle of operation.

The selenium cells used in the study were various sizes of a type of 30-volt cell produced by one manufacturer specifically for magnetic-amplifier applications. It has been found, however, that, although the magnitudes of the characteristics may vary somewhat between different manufacturers and types of cells, the qualitative nature of the characteristics is quite generally that described herein.

General

The experimental technique used in obtaining the data involves the display of the instantaneous voltage-current operating loci on a cathode-ray oscillograph by applying the voltage appearing across the selenium cell to the horizontal axis and a voltage taken from a resistor in series with the cell to the vertical axis, as shown in Fig. 1(A). Quantitative data are then obtained by calibrating the axes and scaling off a photograph of the display.

Fig. 2(A) shows data obtained in this manner for a typical 2-inch by 2-inch 30-volt selenium cell operating at an a-c source frequency of 60 cycles per second (cps) in the circuit of Fig. 1(A). Shown are the operating loop and time base oscillographs of voltage and current oriented so as to correspond to the axes of the loop. Tracing the operation through a cycle, it is seen that, immediately following an instant when the applied voltage is zero,

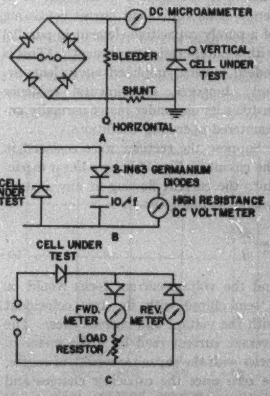


Fig. 1. A—static test circuit. B—peak-reading voltmeter circuit. C—dynamic test circuit

---

Paper 54-164, recommended by the AIEE Metallic Rectifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing December 1, 1953.

G. F. PITTMAN, JR., is with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

the current increases rapidly to a rather high value in the reverse (blocking) direction and then proceeds to decrease as the reverse voltage increases. As the voltage approaches its peak value, the current decreases less rapidly and may increase in magnitude (as shown in the oscillograph). Following passage of the voltage through its peak value, the current drops rather rapidly to zero and reverses direction—an increasingly large current flowing in the forward (conducting) direction of the cell until, just before the voltage becomes zero, a current is flowing in the forward direction approximately equal in magnitude to that observed in the reverse direction immediately following the beginning of the cycle. Upon passage of the voltage through the zero, the current reverses and the cycle repeats. Fig. 2(B) shows similar oscillographs for the same cell operating at the same peak voltage but at an a-c source frequency of 400 cps. The distance *a* shown in both Figs. 2(A) and (B) represents the same current to the scale of each figure. Notice that the 400-cps loop is several times broader and that the current peaks at zero voltage are much larger while the current at the voltage peak remains approximately the same as that observed at 60 cps.

An Equivalent Circuit

A notable feature of the operation observed is that, although the voltage applied to the cell is at all times acting in the reverse direction, a current in the forward direction flows during a large part of the cycle. This behavior is readily explained if the rectifier is thought of in terms of an equivalent circuit consisting of a purely capacitive element in parallel with a purely resistive element. (This is indeed a very simplified equivalent circuit; however, experimental evidence justifies its use under many normally encountered operating conditions.)

Suppose the rectifier were replaced in the circuit of Fig. 1(A) by a linear capacitor; the current flowing at any instant would then be given by

$$i = c\frac{dv}{dt}$$

and the voltage-current locus would be a semi-ellipse with its axes coincident with the voltage and current axes. The average current read by a d-c meter in series with the capacitor would, of course, be zero since the capacitor charges and discharges during each cycle, the current-time integral required being the same for each.

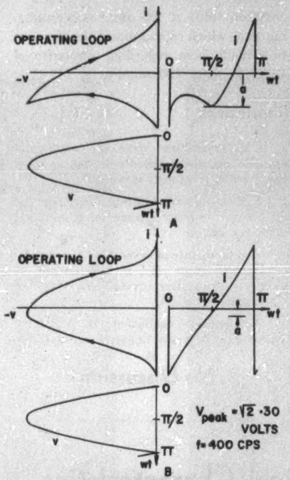


Fig. 2. Instantaneous cell operation with full-wave rectified sinusoidal reverse voltage applied. A—a-c source frequency of 60 cps. B—400 cps

Suppose now a linear resistor is placed in parallel with the capacitor. The straight line voltage-current locus of the resistor must now be added, current-wise to the semi-elliptical locus of the capacitor. The resultant locus, shown in Fig. 3, is a semi-elliptical one skewed downward by an amount depending upon the value of the resistance. The current read by a d-c meter in series with the combination would be the average of the resistive component—the capacitive current again averaging to zero over a cycle. Also the current at the peak voltage point is the peak voltage divided by the resistance, the capacitive current being zero at that instant. The locus of Fig. 3 is seen to exhibit the essential features of that displayed by the selenium rectifier cell, the differences being caused by the nonlinear nature of both the resistance and capacitance exhibited by the selenium cell.

The equivalent circuit parameters of the selenium cell will thus be the resistance

$$r = \frac{v}{i_r}$$

and the differential capacitance

$$c = \frac{i_c}{dv/dt}$$

where

v = applied voltage
i_r = resistive component of cell current
i_c = capacitive component

It may be seen from Fig. 3 that, if both the resistance and the differential capacitance of the cell are single-valued functions of the applied voltage, then, at any given instantaneous voltage, the vertical width of the operating voltage-current loop is twice the capacitive component of current (since dv/dt is a single-valued function of v) while the distance from the voltage axis to the center line of the loop is the resistive component. Thus both parameters may be determined from measurements on the observed loops if dv/dt is assumed known as a function of v from the applied voltage wave form.

The symmetry about the voltage axis of loops observed at frequencies and voltages such that the resistive component of current is very small compared to the capacitive component and the nature of loops observed at very low frequencies where the capacitive component is small indicate that the assumption of single-valued dependencies of the parameters on voltage is a good approximation. A notable exception is found in unformed or very leaky cells where the operating points extend well beyond the knee of the reverse resistance characteristic; in such cells a considerable hysteresis or loop effect is noted in the resistance characteristic itself at the higher values of voltage. This effect is in the opposite direction to that produced by the capacitance of the cell and tends to disappear as the cell forms in under continued application of voltage.

Capacitance

The aforementioned method of determination of differential capacitance yields sufficiently accurate results for voltages in the middle range between zero and the peak value applied; however, it suffers serious inaccuracies near zero voltage

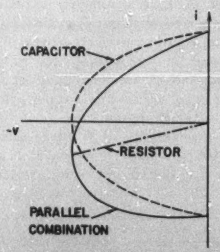


Fig. 3. Instantaneous voltage-current operating loci of a parallel linear resistance-capacitance combination with full-wave rectified sinusoidal voltage applied

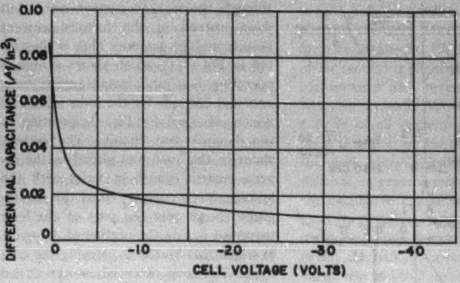


Fig. 4. Differential capacitance of a selenium cell as a function of instantaneous cell voltage

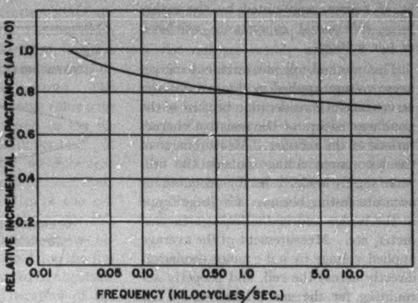


Fig. 5. Relative incremental capacitance of a selenium cell as a function of frequency

since the applied wave form is somewhat distorted in this region due to the capacitive nature of the cell under test. Specifically, the voltage does not fall sharply to zero and increase again, but rather the cusp is rounded off; thus the actual time rate of change of voltage near zero is considerably less than that of a pure rectified sinusoidal wave form. In order to obtain an accurate measure of the capacitance in this region, a small sinusoidal voltage, on the order of 0.1 volt rms, was applied to the cell, superimposed upon a pure direct bias voltage, and the locus of the alternating components of voltage and current observed as before. From the values of the alternating current at the zeros of the alternating voltage, the incremental capacitance of the cell at the direct bias voltage applied was determined. This method was used at zero and near zero bias voltages and also at higher bias voltages to check the correlation between the capacitance determined by this method and that from the previous method. These checks were found to be excellent.

The differential capacitances of a number of samples of various sizes of selenium cells were determined by these methods; typical results are shown in Fig. 4. The differential capacitance, expressed on a unit active area basis, was found to be rather consistent among the cells tested. Note that Fig. 4 extends to slightly positive voltages, the differential capacitance increasing rapidly as the cell voltage goes more positive. No measurements were made at more positive voltages than that shown since the shunting of the capacitance by the very low differential resistance displayed by the cells above their threshold voltage makes such measurements both difficult and insignificant.

The dependence of the differential capacitance upon frequency was also investigated by using the last-mentioned method with frequencies ranging from 20 to 20,000 cps. Typical results of these tests are shown in Fig. 5. The magnitude of the alternating voltage used in the experimentation was found to have no effect upon the value of capacitance determined over a wide range of voltages.

Resistance

The instantaneous reverse resistance characteristic (i_r versus v) of a typical cell, as obtained from the distance from the voltage axis to the center line of the reverse loop, is shown in Fig. 6. This characteristic was found to be essentially independent of the frequency of the rectified sinusoidal voltage applied. This was determined by varying the a-c source frequency and observing the reading of a d-c meter in series with the cell while keeping the peak reverse voltage across the cell constant. Over a frequency range from 60 to 1,000 cps, the current read varied less than 5 per cent. It is to be noted, however, that no good correlation was found between this reverse resistance characteristic of a cell and that measured using pure direct current.

It thus appears that, with respect to its electrical properties, a selenium rectifier cell may be represented, for many purposes, by a nonlinear capacitance in parallel with a nonlinear resistance, the values of both parameters being functions of the instantaneous voltage across the cell. Furthermore, over a very considerable frequency range, these parameters may be considered invariant with the frequency of the voltages to which the cell is subjected. On this basis, and with typical values of the parameters in mind, more careful and accurate consideration may be given operation of selenium rectifiers in the circuits in which they find use. In the following parts of this paper several methods of rectifier testing will be discussed in this light.

The "Static" Test Circuit

This test involves the application of a full-wave rectified sinusoidal voltage to the cell in the reverse direction and the metering of the average current with a d-c meter. This is precisely the circuit of Fig. 1(A) and results in rectifier operation as shown, for example, in Fig. 2. As pointed out previously, the capacitive component of current averages to zero over a cycle, and the meter therefore indicates the average resistive leakage current under the voltage conditions imposed.

The bleeder resistor seen on the d-c side of the full-wave bridge supply of Fig. 1(A) is a very essential part of the test circuit since it provides a path for the forward, discharging capacitive current during the latter part of each cycle. In the absence of this path the voltage wave form across the cell under test would be distorted due to the retarded discharge of the cell capacitance. This situation

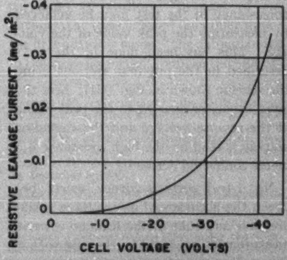


Fig. 6. Instantaneous resistive leakage current of a selenium cell as a function of instantaneous cell voltage

is, of course, aggravated by the rather large differential capacitance exhibited at low voltages.

The method of measurement of reverse voltage applied in the test also deserves careful consideration because of the nonlinear nature of the resistive characteristic of the rectifier. Measurement of the a-c source voltage outside the full-wave supply bridge is hardly adequate for accurate testing because of voltage drops in the bridge and in the leakage current meter, etc. Measurement of the average applied voltage by a d-c meter connected directly across the cell, and properly accounting for the current drawn by the meter, may be adequate provided little distortion is present in the voltage wave form due to the cell capacitance and impedances in series with the cell. The presence of sizable series resistances in the circuit can result in serious inconsistencies in the test results. Due to the nature of the nonlinearity of the reverse resistance characteristic, the major contribution to the average resistive leakage current measured is from the region of operation at and near the peak reverse voltage; it is essential, therefore, that the same peak value of voltage be applied to all cells if comparable average resistive leakage currents are to be measured. The relatively large current flowing in the circuit during the periods of high instantaneous reverse voltages, especially if a poor cell is under test, may because considerable voltage drops to appear across such series resistances, resulting in the voltage wave form applied to the cell being flat-topped. Thus, if the average voltage is maintained the same for all cells, leakier cells will be subjected to lower peak voltages than less leaky ones, and will thus display lower values of resistive leakage current than they would in a truly comparable test.

Because of the sensitivity of the average resistive leakage current to the peak value of applied voltage, a greater degree of consistency in the test may be achieved by measuring the peak value of the voltage. This has been done in the tests described in the previous sections using the scheme shown in Fig. 1(B); here the large capacitor charges to the peak value of the reverse voltage and the capacitor voltage is read by the high resistance d-c meter connected across it.

No 1-to-1 correspondence exists between the accuracy of any voltage measurement and that of the leakage current measurement as would be the case with a linear element under test; this is particularly true when the differential reverse resistance (dv/di_r) attains a relatively low value at the peak voltage applied in the tests. In several tests of poorer cells a few per-cent change in the peak reverse voltage produced a 100-per-cent change in the average resistive leakage current measured.

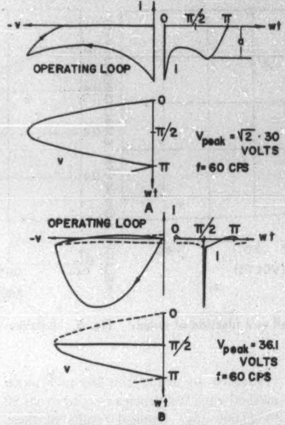


Fig. 7. A—Instantaneous cell operation in the dynamic test circuit. B—Instantaneous cell operation with typical magnetic-amplifier applied voltage wave form

Summarizing, care must be exercised in conducting this type of test to insure the application of undistorted rectified sinusoidal voltage to all cells tested, and the voltage applied must be accurately determined if consistent results are to be attained. The test yields the average resistive leakage current under the voltage conditions imposed but it gives no information as to the instantaneous values of the total current passed by the cell.

The "Dynamic" Test Circuit

In this test the cell is operated as a half-wave rectifier feeding a resistive load, as shown in Fig. 1(C). The current passed by the cell is split into two branches, all current in the forward direction of the cell flowing in one branch and all current in the reverse direction flowing in the other. The average forward current and the average dynamic leakage current are measured by d-c meters in the respective branches.

In this circuit the current which the cell passes in the reverse half-cycle after the capacitive component causes the total current to go positive does not flow through the leakage current meter but flows instead in the forward current branch and its meter. This is illustrated in Fig. 7(A) which shows a reverse operating loop and instantaneous voltages and currents for the same 2-inch by 2-inch cell used for Fig. 2 operating in the dynamic test circuit. The current shunt in this case was placed in the reverse current branch in series with the leakage current meter; thus the oscillograph shows only the part of the loop traversed during the portion of the cycle in which this branch is carrying the current. The loop observed is seen to be essentially the portion below the voltage axis of the loop observed in the static test; the current trace shows the duration of current flow in the reverse current branch to be only slightly greater than a quarter cycle, as explained before.

The switching of the current path results in a net contribution to the average dynamic leakage current by the capacitive component since the leakage meter sees the entire charging current-time integral of the cell capacitance but only a very small portion of the discharging current-time integral (if the resistive leakage component were zero, the current read by the leakage meter would actually be the charging current alone averaged over a cycle). For typical cells tested the capacitive contribution to the average dynamic leakage current measured at 60 cps is of the order of three to four times the resistive contribution as deduced from static tests, depending on the peak reverse voltage applied; at 400 cps the ratio may be as high as 15 to 1 if reduced voltage is used in the tests.

The presence of the large capacitive component in the current measured results in a strong dependence of the average dynamic leakage current upon the a-c frequency used; in fact, if the resistive component were negligible compared to the capacitive component, which is approached in the case of good cells operating at reduced voltages, the leakage measured by this method would increase linearly with frequency.

The same problems regarding the measurement of the reverse voltage applied to the cell are encountered in this test as in the static test. The sensitivity of the leakage measurement to the peak applied voltage is reduced in this case, however, because of the addition of the large capacitive component of current to the sensitive resistive component. Distortion of the reverse voltage wave form applied in this test due to slow discharging of the cell capacitance may result if very high values of load resistance are used (or if the forward current branch is opened). Care must also be exercised, of course, to insure that the rectifying elements used in the two branches have sufficiently low leakages that the measurement is not affected.

The measurement is, however, unavoidably influenced by the capacitance of the rectifier used in the forward current branch. It must carry the forward current of the cell under test and must therefore be of comparable size. Thus, as the alternating voltage passes through zero going negative the forward current rectifier presents a rather large capacitance shunting a portion of the high negative charging current taken by the cell under test at this time around the reverse current branch. This results in the reverse current meter reading less than the true average value of all current passed in the negative direction by the cell under test. The magnitude of this effect is of course proportional to the capacitance of the rectifier used in the forward current branch, hence, the size of its cells. The measured average dynamic leakage is thus dependent upon the cell size used in the forward current rectifier even though the resistive leakage may be low enough to have an insignificant effect upon the measurement; this has been noted in the production testing of cells by means of this circuit. One solution is to use in all cases the same cell size in the forward current branch as in the rectifier under test; this should at least give a consistent test within each cell size.

Summarizing, the leakage measurement of the dynamic test is less sensitive to variations in the peak reverse voltage applied to the cell; however, consideration must be given the effects upon the test of the rectifying elements used in separating forward and reverse current. The test yields, ideally, the average value over a cycle of all current passed in the reverse direction by the cell under test. It therefore gives a leakage measurement which is, in general, more dependent upon the capacitive characteristics of the cell than its resistive characteristic; for this reason no direct correlation exists between the leakage measured by the dynamic test and that measured by the static test, even though the operation of the tested rectifier in the reverse direction is essentially the same in both cases.

Operation with a Typical Magnetic-Amplifier Reverse Voltage Wave Form

To illustrate the effect upon the operation of the cell of reverse voltage wave forms which may be encountered in magnetic-amplifier applications, the data shown in solid line in Fig. 7(B) were obtained by applying to the cell the output voltage of a bridge-type magnetic amplifier. Note particularly the large spike of reverse current and consequent wide swing of the operating loop during the steep rise of applied voltage, indicating the relatively low impedance exhibited by the cell during this brief interval. Shown in dashed line for comparison are data for the same cell with a full-wave rectified sinusoidal applied voltage with the same peak value. In this example the significance of the applied reverse voltage wave form and the capacitive properties of the cell are readily apparent. It is expected that a recognition of the full significance of the instantaneous reverse characteristics of selenium rectifiers as they affect magnetic-amplifier performance will be made possible by an awareness of these characteristics in conjunction with the continually improving understanding of the operating mechanism of magnetic-amplifier circuits.

Conclusions

The electrical properties of selenium rectifiers in the reverse direction may be conveniently represented by an equivalent parallel combination of a nonlinear resistor and a nonlinear capacitor for many engineering purposes. Both the resistance and the differential capacitance show a decreasing trend with increasing reverse voltage while, to a good approximation, both parameters may be considered independent of the frequency of applied voltage.

In critical applications, the existence of both resistive and capacitive effects must be recognized and their magnitudes and variations considered if a realistic understanding of the rectifier operation and its effect upon the remainder of the circuit is to be had. This applies equally well to rectifier test circuits since the leakage quantity measured may be greatly influenced by these effects.

◆

No Discussion

Bibliography on Data Storage and Recording

GERHARD L. HOLLANDER
ASSOCIATE MEMBER AIEE

THIS bibliography contains 330 titles and abstracts on publications applicable to the fields of data storage, recording, analogue-to-digital conversion, data presentation, and telemetering.

It contains references pertaining to topics connected with storage of information in a minimum of space. In the process of storing and recording data, various factors must be considered, such as: the characteristics of the storage media themselves; methods of data representation including optimum coding, modulation methods, and analogue-to-digital conversion; and the presentation of data at read-out. Thus, some references may prove useful to other fields with similar problems.

An exhaustive literature search for the period of 1948 to the middle of 1952 provides the basis for this compilation. Many earlier and later references are also included.

Most references contain a short note describing the content of the article as it pertains to data storage. The detail or length of these abstracts does not indicate the value of a publication.

Entries are arranged alphabetically by surnames of authors. In case of coauthors, the item is listed only once under the first author's name. The subject index to the more important items in the bibliog-

---

Paper 54-169, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted March 9, 1953; made available for printing December 8, 1953.

Gerhard L. Hollander is with the Massachusetts Institute of Technology, Cambridge, Mass.

This work was supported in part by the Bureau of Ordnance Contract *NOrd-11799*, under the technical direction of the United States Naval Underwater Ordnance Station, for the purpose of recording data on the performance of ranged torpedoes. The author gratefully acknowledges the aid of Miss Betty Jeane Spear in the preparation of the final manuscript.

raphy should be consulted before reading the bibliography. A list of all periodicals which have been cited and their place of publication is given on page 58.

1. AUTOMATIC ACCOUNTING MACHINE. *Engineering*, vol. 171, April 27, 1951, p. 516. Announcement of new National Cash Register Company punch card model. No technical information.

2. AUTOMATIC DIGITAL COMPUTER AT MANCHESTER UNIVERSITY. *Engineer*, vol. 192, July 1951, pp. 47–49, 91–93, 126–28. 1. General description of Ferranti *MK 1* computer. Uses Williams tubes and magnetic drum memories. 2. Gate circuits, arithmetic circuits, and some control circuits. 3. Equipment description.

3. COMPUTERS. *Chemical Engineering*, vol. 57, Dec. 1950, pp. 117–30. Survey of various high-speed computers. Features of digital and analogue computers. Use of punched card machinery. Numerical systems. Bibliography.

4. HIGH SPEED DATA RECORDER. *Electrical Manufacturing*, vol. 50, Oct. 1952, pp. 162, 164. One-page description of Anderson-Nichols printer which has a rate of 15 lines per second.

5. PHOTOGRAPHIC MEMORY DEVICE. *Mechanical Engineering*, vol. 70, Jan. 1948, p. 25. Short announcement of Eastman Kodak's device for project Whirlwind. Summary of memory characteristics.

6. RECORDER SPECIFICATION. *Tele-Tech*, vol. 9, April 1950, pp. 34–36. Three-page tabulation of characteristics of home, commercial, and broadcast equipment, both disk and magnetic tape. Audio point of view.

7. MULTICHANNEL ANALOG INPUT-OUTPUT CONVERSION SYSTEM FOR DIGITAL COMPUTER, **P. A. Adamson, M. L. MacKnight**. *Proceedings*, Institute of Radio Engineers, vol. 41, March 1953, p. 405. Two-paragraph abstract of reference 188.

8. FERRITE DEVELOPMENT—QUARTERLY PROGRESS REPORTS, **Dr. E. Albers-Schoenberg et al.** *Contract DA 36-039-sc-5449*, General Ceramics and Steatite Corporation, Keasby, N. J. Progress reports containing some technical data:

   1. *Report no. 1*, from 4-1-51 to 6-30-51, 15 pages: Magnetization curves and temperature effects on initial permeability for various new ferrites.

   2. *Report no. 8*, from 1-1-51 to 3-31-51, 26 pages: Index of tables and curves in first eight reports. Magnetization hysteresis, permeability, and firing angle curves.

   3. *Report no. 4*, from 1-1-52 to 3-21-52, 13 pages: Further work on ferrites. Graphs of Q as function of other variables.

9. A BIBLIOGRAPHY OF MAGNETIC RECORDING, **Donald W. Aldous**. *Electronic Engineering*, vol. 19, Dec. 1947, pp. 390–91. References without abstracts to 96 papers and 10 patents.

Subject Index

CIRCUITS

Comparison circuits 111, 188, 215, 269
Counters 71, 78, 266, 313
Oscillators 171
Sweep circuits 269
Switching circuits 10, 23, 24, 39, 40, 50, 78, 141, 166, 222, 307

CODING SYSTEMS 18, 47, 76, 78, 127, 179, 250, 313

Error-checking codes 103, 118, 133, 134

COMPUTERS, DESCRIPTION OF DIGITAL 2, 12, 20, 28, 72, 78, 122, 139, 212, 300, 315

DATA PRESENTATION 78

Curves (analogue)
  Curve printer 24, 29, 168
  Oscillograph 228, 231, 254, 276, 294, 312
  Oscilloscope 287
Numerical (digital)
  Paper (printer) 4, 35, 248
  Oscilloscope 94

DATA REDUCTION EQUIPMENT 37, 65, 232, 248

DATA STORAGE

General storage investigations 76, 78, 80, 106, 133, 134, 173
  Library catalog storage 18
  Three-dimensional storage 88, 89
Storage methods; see also Circuits, Transistors
  Comparison of storage methods 78, 106, 134, 139, 173, 174
  Electrostatic storage tubes 2, 63, 78, 145, 155, 157, 173, 174, 236, 323
  Ferroelectric storage 11, 40, 233
    Ferrite characteristics 8
  Glow tubes 20, 88, 193
  Magnetic cores 89, 105, 153, 201, 203, 204, 205, 227, 289, 292, 305, 308
  Magnetic (surface) recording 128
    Bibliographies 9, 301, 324
    Boundary-displacement recording 61
    Equipment design 22, 52, 133, 134, 251
    Head design 162, 183, 184, 239
      Static reading head 49, 56, 115, 133, 134, 244, 256, 319
    Recorders,
      Drum recorders 2, 30, 78, 321
      Tape recorders 6, 32, 41, 84, 133, 134, 152, 189
    Surface investigations 136, 148, 158, 160, 286, 290
    Test results, 57, 62, 101, 124, 136, 147, 158, 167, 182, 211, 239, 240, 241, 243, 244, 278, 303, 316, 317
    Theory 17, 22, 31, 96, 112, 124, 136, 137, 159, 161, 182, 206, 211, 303, 316, 317, 318, 326, 330
  Mechanisms 25
  Molecular disturbances 116
  Phosphor 207, 298
  Photographic storage 78, 208, 253, 298, 299
    Color film 172
    Devices 5, 19, 68, 75, 172, 217, 242, 255, 260, 282, 297
    Film characteristics 69, 70, 133, 134, 149, 163, 164, 186, 198, 260, 264, 311
    Light source characteristics 90, 92, 102, 113, 114, 247, 309, 325
    Playback 221
    Photographic theory 91, 119, 133, 134, 197, 198
  Scribed surfaces 14, 150
    Soot 73
  Sonic delay lines 320
  Voltage or heat-sensitive paper 29, 35, 142, 181, 190, 218, 276

INFORMATION THEORY 36, 51, 83, 95, 165, 179, 192, 194, 216, 223, 230, 238, 261, 262, 263, 283, 295, 296, 302

INSTRUMENTATION 74, 81, 225, 328

MODULATION (DATA REPRESENTATION)

Comparison of modulation methods 81, 125, 126, 214, 224, 235
Amplitude modulation 281
Boundary-displacement modulation 61
Frequency modulation,
  Demodulators 13, 16, 133, 134
  Errors and noise 52, 170
  Recording equipment 32, 84, 213
Incremental representation 19
  Commutators 15
Phase-shift modulation 120, 146
Pulse-code modulation 27, 48, 64, 85, 104, 109, 110, 196, 216, 219, 245, 246, 272, 275
  Analogue-to-digital conversion 78
    Comparison of converters (surveys) 43, 44, 45, 132, 176, 177, 279
    Converter applications 27, 35, 85
    Principles of converters,
      Voltage converters 21, 104, 111, 121, 188, 215, 219, 245, 248, 258, 259, 266, 267, 268, 269, 271, 272, 273, 275
      Position converters 87, 175, 252, 279, 280, 329
    Digital-to-analogue conversion 64, 246, 249, 268, 284, 304, 327
Pulse-time modulation 26, 100, 199

TELEMETERING

Circuits 171
Data reduction 37, 65, 248
Modulation methods 170, 179, 180, 214, 281
Systems 26, 42, 120, 125, 126, 195, 199
Use of recording 32, 33, 213

TRANSDUCERS 54, 142

TRANSISTORS 23, 237

Counters 71, 138, 293
Oscillator 171
Switching circuits 10, 23, 138
Theory 23

10. Transistors in Switching Circuits, **A. Eugene Anderson.** *Bell System Technical Journal*, vol. 31, Nov. 1952, pp. 1207–49.

11. Ferroelectric Storage Elements for Digital Computers and Switching Systems, **J. R. Anderson.** *Electrical Engineering*, vol. 71, Oct. 1952, pp. 916–22.

12. The Bell Computer, Model VI, **E. G. Andrews.** *Electrical Engineering*, vol. 68, Sept. 1949, pp. 751–56. Computer description. Storage is on relays.

13. A Transcriber for Magnetically Recorded F-M Signals, **Ernest J. Angelo, Jr.** *Report 68*, Massachusetts Institute of Technology, Cambridge, Mass., Jan. 1951, 22 pp. Design equations and circuits for a frequency-modulated demodulator which counts zero-crossings. Tape speed variations are corrected by comparison with an unmodulated reference channel. Accuracy on the order of 0.1 per cent.

14. New Danish Playback Machine, **Svend Anker-Rasmussen.** *Audio Engineering*, vol. 35, Sept. 1951, pp. 40, 43. Brief description of record player using a polyvinyl-chloride tape having 4 tracks per millimeter. Intended to replace long-playing records.

15. Development of Mechanical Devices for Multichannel Sampling and Automatic Calibration in Missile Telemetering. *Report JHU/APL/CM-543*, Applied Science Corporation of Princeton, Princeton, N. J., 1949, 53 pp. Mechanical design information and test data on low-speed (1,800 rpm) and high-speed (24,000 rpm) commutators having life of 100 hours and 30 hours respectively.

16. The Statistical Properties of the Output of a Frequency Sensitive Device, **George R. Arthur.** *Convention Record*, Institute of Radio Engineers, pt. 8, 1953, pp. 81–90. Includes information applicable to frequency-modulated discriminator analysis.

17. An investigation into the Mechanism of Magnetic Tape Recording, **P. E. Axon.** *Paper 1149*, Institution of Electrical Engineers, London, England, vol. 99, pt. III, May 1952, pp. 109-26. Property of recording with variation of bias, etc.

18. Electronic Digital Machines for High-Speed Information Searching, **Philip R. Bagley.** *M.S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., Aug. 1951, 133 pp. Survey of means to store index of entire library suitable for automatic searching. Punched card and digital electronic machines compared. Various coding systems compared. Existing systems described. With magnetic tape author expects that 5,000,000 documents at 900 bits per document can be scanned in 1 hour. Bibliography.

19. Digital Recording Techniques and Recorders, **Saul Baker.** *Report NAVORD 1214*, United States Naval Ordnance Test Station, Inyokern, Calif., Aug. 4, 1950, 35 pp. Detailed explanation and description of devices to quantize and record shaft movements of measuring instruments or transducers by commutator technique. After comparing various methods of storing data, author decides to record incremental data. Final devices capable of withstanding 200–600g occupy about 300 cubic inches. Neon tubes record commutator data on film.

20. An Electronic Digital Computer Using Cold Cathode Counting Tubes for Storage I, II. **R. C. M. Barnes, Cooke-Yarborough, D. G. A. Thomas.** *Electronic Engineering*, vol. 23, Aug.-Sept. 1951,pp. 286–91, 341–43. Over-all description of the features and operation of a simple, low-speed computer. Storage elements are dekatrons.

21. The Binary Quantizer, **Kay Howard Barney**, *Electrical Engineering*, vol. 68, Nov. 1949, pp. 962–67. Counter circuit which counts backward and forward. Counter puts out analogue voltage also which is compared to input, and error signal determines whether forward or backward count is needed to reduce error. Circuit also computes approximate derivative of input. Good explanation with block diagrams and detailed circuit description. Experimental circuit operated at 4 microseconds.

22. Magnetic Recording (book), **S. J. Begun.** Murray Hill Books, New York, N. Y., 1949.

23. The Transistor. Bell Telephone Laboratories, Inc., New York, N. Y., 1951, 792 pp. Symposium of 35 papers on the characteristics and applications of transistors.

24. The Teleplotter, A Digital Plotting Device, **Donald E. Belloff.** *Proceedings*, Electronic Computer Symposium. May 1952, pp. XVIII: 1–7. Description of a plotting device which accepts rectangular coordinates in digital form and plots the point by counting lines on the graph paper. Speeds up to 75 points per minute and accuracies of 0.01 per cent are possible. Any graph paper can be used. Description of a few novel circuits including a thyratron acting simultaneously as switch and memory.

25. Electromechanical Registers as Used in Radioactive Counting Systems, **A. E. Bennett.** *Electronic Engineering*, vol. 23, March 1951, pp. 81–85. Two new electromechanical registers: 1. Escapement type—maximum speed 70 steps per second; 2. Motor-driven type with magnetic clutch—maximum speed 60 steps per second.

26. Matrix Telemetering System, **Nolan R. Best.** *Electronics*, vol. 23, Aug. 1950, pp. 82–85. Description and block diagrams of 30-channel pulse-time modulated, sequential telemetering equipment with a 1-per-cent accuracy and 312.5-cycle repetition rate.

27. Pulse Code Modulation, **H. S. Black, J. O. Edson.** *AIEE Transactions*, vol. 66, 1947, pp. 895–99. Description and block diagrams of pulse code modulation transmitter and receiver system. Eight channels are sampled 8,000 times per second each with 5-binary-digit accuracy (320 kc). Clock pulse coder is described briefly.

28. The Logical Design of the Raytheon Computer, **Richard M. Bloch, et al.** *Mathematical Tables and Other Aids to Computation*, vol. 3, Oct. 1948, pp. 286–95. Description of computer including magnetic tape memory. No details on memory.

29. A High-Speed Curve Printer, **Alvin Victor Block.** *M.S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1950, 50 pp. Design and circuits of a curve printer using Alfax paper. A motor-driven helix acts as one electrode. Accuracy ±1 per cent. Rate 100 samples per second.

30. Magnetic Digital Storage System, **A. D. Booth.** *Electronic Engineering*, vol. 21, July 1949, pp. 234–38. Rotating drum system.

31. On Two Problems in Potential Theory and Their Application to the Design of Magnetic Recording Heads for Digital Computers, **A. D. Booth.** *British Journal of Applied Physics*, vol. 3, Oct. 1952, pp. 307–08. Two sets of examples show that the external field from the recording head gap is very insensitive to the pole-piece shape. No experimental verifications.

32. Tape Recording for Telemetering and Data Analysis, **Kenneth B. Boothe.** *Tele-Tech*, vol. 11, May 1952, pp. 44–46, 116; vol. 11, June 1952, pp. 56–57, 90, 92, 94. Description and discussion of the Ampex models 303, 306, 307, 309, and 500 tape recorders.

33. Uses of Magnetic Tape Recording in Telemetering, **Kenneth B. Boothe.** *Instruments*, vol. 23, Nov. 1950, pp. 1186–88. Two-page paper describing some applications of magnetic tape recorders with frequent reference to Ampex model *302*. Data and limitations of magnetic recording given.

34. Principles and Methods of Telemetering (book), **P. A. Borden, G. M. Thynell.** Reinhold Publishing Corporation, New York, N. Y., 1948. Radio telemetering not considered.

35. High-Speed Recording Potentiometer, **Ralph B. Bowersox, C. G. Hylkema.** *CIT/JPL Memorandum 20-69*, California Institute of Technology, Pasadena, Calif., Jan. 16, 1952, 25 pp. Detailed description and circuit diagrams for recorder having a rate of 20 samples per second, an accuracy of ±0.3 per cent, and a precision of 0.2 per cent. Analogue voltage is coded in an analogue-to-digital converter employing clock pulses and sweep timing. Clock-pulse rate 27.5 kc. Digits are recorded on 6-inch electrosensitive chart paper.

36. Information Theory and Most Efficient Codings for Communication and Memory Devices, **L. Brillouin.** *Journal of Applied Physics*, vol. 22, Sept. 1951, pp. 1108–11. Shannon's channel capacity theorem used to find most efficient code as that yielding most probable distribution of code symbols. Specific rule for most probable distribution. References to communication channels and computer memories.

37. Data Reduction Methods for Radio Telemetering Systems, **J. F. Brinster, E. Donath.** Applied Science Corporation of Princeton, Princeton, N. J., 1952, 13 pp. Paper presented at Western Electronic Show and Convention. Functional description of MADAM, an automatic data reduction equipment.

38. Instrumentation for Flight Testing Airplanes, **W. G. Brombacher.** *Instruments*, vol. 20, Aug. 1947, pp. 700–08. Various instruments used for flight testing. Many tables classifying instruments. Recording is accomplished by photographing the instrument panels. Bibliography on instrumentation.

39. Rectifier Network for Multiposition Switching, **David R. Brown, Nathaniel Rochester.** *Proceedings*, Institute of Radio Engineers, vol. 37, Feb. 1949, pp. 139–47. Crystal matrices are described. Design of matrices using a minimum number of elements is illustrated. Appendix gives mathematical derivations for minimum element switching methods.

40. Ferroelectrics for Digital Information Storage and Switching, **Dudley Allen Buck.** *Whirlwind Report R-212*, Massachusetts Institute of Technology, Cambridge, Mass., June 5, 1952, 53 pp. Also *M.S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., May 16, 1952. Description and circuits of pulse test equipment for ferroelectrics. Test procedures and results of temperature effects on switching time and switching loss. Ferroelectric memory matrix. Choice of selection ratio. Tests on memory matrix. Ferroelectric switch.

41. The Multi Channel Recorder, **W. C. Buhler, et al.** *Technical Data Digest*, United States Air Force, vol. 13, Aug. 15, 1948, pp. 11–15. Description of a 200-pound gondola with instruments and 7-channel recorder for parachute research. Withstands 75g shock.

42. Ein Messwertumformer nach dem Frequenzverfahren, **Rudolph Burgholz.** *Archiv für technisches Messen* no. 174, July 1950, T84, 2 pp. (Reference No. J86-1). Short description of an industrial telemetering system which transmits data sequentially in different audio-frequency channels.

43. Some Techniques of Analog-to-Digital Conversion, **Harry E. Burke, Jr.** *Proceedings*, Electronic Computer Symposium, May 1952, pp. XVII: 1–8. Elementary classification of various analogue-to-digital conversion methods.

44. A Survey of Analog-to-Digital Converters, **Harry E. Burke, Jr.** *Proceedings*, Institute of Radio Engineers, vol. 41, Oct. 1953, pp. 1455–62. Description of various quantizers. Bibliography including some unpublished reports.

45. Review of Input and Output Equipment Used in Computing Systems. *AIEE Special Publication S-53*, "A Survey of Analogue-to-Digital Converters," **Harry E. Burke, Jr.**, 1953, pp. 98–105. Identical to reference 44.

46. Survey of Electronic Commutation Methods, **Robert S. Butts.** *Proceedings*, Institute of Radio Engineers, vol. 39, March 1951, p. 296. Brief abstract of convention paper.

47. Proposed Digital Method for Recording Shaft Position, **W. M. Cady.** *Technical Memorandum 759-050950*, United States Naval Ordnance Test Station, Inyokern, Calif., July 19, 1950, 29 pp. Exhaustive study of efficient coding using several intensity levels and contextual meanings. This is applicable for coding mechanical motions varying in single steps. Coding and inverse tables given for context of 2, 3, 4, 6, and 12. Circuit diagrams for operation of lamp circuits. Mathematical expressions for maximum number of states using various level and context combinations.

48. Decoding in Pulse Code Modulation, **R. L. Carbrey.** *Bell Laboratories Record*, vol. 26, Nov. 1948, pp. 451–55.

49. A Sensitive Magnetometer for Very Small Areas, **Daryl M. Chapin.** *Review of Scientific Instruments*, vol. 20, Dec. 1949, pp. 945–46. Also *Bell Telephone System Monograph 1720*, Bell Laboratories, Inc., New York, N. Y., April 1950, 2 pp. Description and circuit for a quartz-driven loop probe, measuring an area less than 1 millimeter by 0.15 millimeter and sensitive to less than 1 gauss.

50. Diode Coincidence and Mixing Circuits in Digital Computers, **Tung Chang Chen.** *Proceedings*, Institute of Radio Engineers, vol. 38, May 1950, pp. 511–14. Coincidence ("and") and mixing ("or") circuits using "ideal" diodes are analyzed. Rise times and input impedances are considered.

51. History of Theory of Information, **E. C. Cherry.** *Proceedings*, Institution of Electrical Engineers, vol. 98, pt. III, Sept. 1951, pp. 383–93.

52. Errors Introduced into F-M Carrier Measuring Systems, **James C. Coe.** *Instruments*, vol. 24, March 1951, pp. 274–78. Time-base variations account for errors in frequency-modulated carrier systems if recording is done before demodulation. Experimental speed characteristics of various constant-speed motors. Accuracy of 1 per cent was desired. Relation of scale factor versus per cent error caused by time-base variation is given.

53. Magnetic Drum Storage for Digital Information Processing Systems, **Arnold A. Cohen.** *Mathematical Tables and Other Aids to Computation*, vol. 4, Jan. 1950, pp. 31–39. Elementary description and operating characteristics of the Engineering Research Associates storage drum system.

54. Strain-Gage Dynamometers and Indicators, **Richard E. Conover.** *Instruments*, vol. 23, May 1950, pp. 445–48. Methods to reduce errors in strain gauges used in air-borne recording.

55. Counting of Random Pulses, **E. H. Cooke-Yarborough.** *Journal*, British Institution of Radio Engineers, vol. 11, Sept. 1951, pp. 367–80.

56. Status of Research on Magnetic Recording, **Gerald Cooper.** *Whirlwind Engineering Note E-195*, Massachusetts Institute of Technology, Cambridge, Mass., Feb. 14, 1949, 8 pp. Describes test equipment and design of static reading head.

57. Pulse Packing in Magnetic Recording Wire, **I. L. Cooter.** *Journal of Research*, United States Bureau of Standards, vol. 44, Feb. 1950, pp. 163–72. Method and apparatus are described for determining relative pulse packing of different recording wire. Typical curves show influence of amplitude, duration, and repetition rate of magnetizing pulses on the pulse packing.

58. A Digital Automatic Multiple Pressure Recorder, **Bert A. Coss et al.** *Technical Note 2880.* National Advisory Committee for Aeronautics, Cleveland, Ohio, Jan. 1953, 24 pp. A large number of pressures are measured and recorded by scanning all of the unknown pressures by a linear pressure sweep, detecting and stopping a counter at coincidence. Sweep rate about 2.5 minutes.

59. Nouvelle Méthode Pour Accroître la Sensibilité à la Tension des Galvanomètres à Cadre Mobile, **Jean Coursaget.** *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, vol. 227, Oct. 4, 1948, pp. 673–75. This describes a new method of increasing the sensitivity of the torque in moving-coil galvanometers. A unit employing light beam, photocell, and servo is suggested.

60. Advanced Engineering Research Study of High Ffrequency Communication Methods, **Murray G. Crosby.** Crosby Laboratories, Mineola, N. Y., March 31, 1950 (final report), 50 pp. Study of sky-wave propagation.

61. Boundary-Displacement Magnetic Recording, **H. L. Daniels.** *Electronics*, vol. 25, April 1952, pp. 116–20. Description of boundary-displacement recording. Since almost the entire recording medium is saturated at all times, the nonlinearities of the medium do not affect the output linearity. Noise is reduced and bias level is not critical. Zero-signal noise is comparable to d-c saturation erasing, but dynamic range approaches that obtained by high-quality supersonic bias.

62. The Use of Magnetic Wire Recording in the FM-FM Telemetering System, **Fred H. Davis, Frank W. Lehan.** *CIT/JPL Memorandum 4-44*, California Institute of Technology, Pasadena, Calif., Sept. 27, 1948, 11 pp. A wire recorder was used to monitor frequency-modulated–frequency-modulated system by recording signal before separating subcarriers. Curves show frequency response of Magnecorder *SD-1F* with parameters of head cleanliness, fading due to wire twist, and tape speed.

63. Electrostatic Storage Tube, **S. H. Dodd, H. Klemperer, P. Youtz.** *Electrical Engineering*, vol. 69, Nov. 1950, pp. 990–95.

64. Magnetic Recording of Sound by Pulse Coding, **Ladislav Dolansky.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., May 1949. Design of playback apparatus for 7-digit pulse-coded audio up to 25 kc. Three different decoders described. Elementary discussion of magnetic recording and its errors. Bibliography.

65. Automatic Data Reduction, **Erwin Donath.** Applied Science Corporation of Princeton, Princeton, N. J., Dec. 1951, 10 pp. Description and block diagrams of MADAM, automatic telemetering data reduction equipment.

66. New Electromechanical System Provides Fast Access to Punched Card Data Files, **J. H. Drillick.** *Product Engineering*, vol. 22, Oct. 1951, pp. 176–78. Punched card random access memory with high storage capacity.

67. Elektrische Fernuebertragungs-und Rechenschaltungen (in English), **Drodofsky.** *Report F-TS-3092RE*, Air Materiel Command, Wright Field, Dayton, Ohio, Jan. 1948, 5 pp. Electric circuits for telemetering and computing. Basic Wheatstone bridge circuits and their accuracy considerations. Telemetering example.

68. The Kodak Digital Reader-Recorder, Model A. Eastman Kodak Company, Rochester, N. Y., June 1950, 20 pp. Equipment instruction book primarily for operator and maintenance man. 25 digits per line in "0" and "1" representation, or 50 bits per line. Maximum reading rate 2,000 lines per second. There are 100 lines per inch.

69. Kodak Films. Eastman Kodak Company, Rochester, N. Y., 1951, 5th ed., 64 pages. (Kodak Data Book Series) Characteristics of Kodak films and an introductory. section explaining their significance.

70. Kodak Linagraph Papers and Films for Instrument Recording. Eastman Kodak Company, Rochester, N. Y., April 1952, 4 pp. Contains *H* and *D* curves, and spectral sensitivities of six papers and five films.

71. Counter Circuits Using Transistors, **E. Eberhard, R. O. Endres, R. P. Moore.** *RCA Review*, vol. 10, Dec. 1949, pp. 459–76. Excellent paper on various transistor circuits. One-shot and free-running multivibrators, counting circuits, etc. Shows bistable stage using single transistor. Gives circuit constants and circuit operating limitations: speed, stability with supply voltage, trigger voltage, temperature, and age variation. Design information for all circuits and design equations for single transistor, bistable circuit.

72. The IBM Pluggable Sequence Relay Calculator, **W. J. Eckert.** *Mathematical Tables and Other Aids to Computation*, vol. 3, July 1948, pp. 149–61. Machine description including relay storage registers and card punching facility.

73. High Resolution Recording with Soot, **K. R. Eldredge.** *Review of Scientific Instruments*, vol. 21, March 1950, pp. 199–202. Soot coating with hydrocarbon solvent permits a resolution of 500 lines per centimeter with a recording force of 1 milligram, and a coefficient of friction less than one. Line width about 10 microns. Article describes preparation of soot surface and the stylus requirements.

74. Automatic Flight-Research Instrumentation of the Single-Seater Fighter, **S. B. Elggren.** *Instruments*, vol. 18, Jan. 1945, pp. 2–9, 44. Primarily a paper on instrumentation. Airborne recording.

75. A Sensitive Photographic Recorder, **A. Elliott, E. J. Ambrose.** *Journal of Scientific Instruments*, vol. 24, Dec. 1947, pp. 324–25. Photographic recording of micrometer deflection.

76. Storage and Switching, **A. G. Emslie.** *Quarterly Progress Report 1*, Williams College, Williamstown, Mass., April 11, 1950, 11 pp. Survey of storage devices. Analogue and digital representation are compared. Examples of storage and channel capacity use ϵ (2.718) as base for unit of information.

77. Storage and Switching, **A. G. Emslie.** *Progress Report 2*, Williams College, Willamstown, Mass., Sept. 12, 1950, 20 pp. Comparison of storage media capable of storing 10^4 words having 100 digits per word with 1-microsecond access time. Volume must be limited. After considering multivibrators, Williams tubes, and transistors, author concludes that goal may be reached with tube having a metallic mosaic storage surface.

78. High-Speed Computing Devices (book), **Engineering Research Associates Staff.** McGraw-Hill Book Company, New York, N. Y., 1950. A comprehensive treatment of the basic components of computers and their application. Fundamentals of machine computation and numerical analysis are given. Chapter 14 surveys storage media; chapter 15, data converters. Bibliography.

79. Summary of Characteristics, Magnetic Drum Binary Computer. *Report PX 29201, Contract CST-10133*, Engineering Research Associates, Inc., Saint Paul, Minn., Nov. 30, 1948, 36 pp. A proposal for a new computer describing the functional components, computing characteristics including order codes and operating time, and some maintenance information.

80. The Storage of Information, **P. M. Erlandson.** *Report JHU/APL/CF-1339*, University of Texas, Austin, Tex., Oct. 1, 1949, 17 pp. Comparison of data storage to information transmission process. Very general, philosophical treatment.

81. Telemetering and Direct Frequency Measurement, **Paul M. Erlandson.** *Tele-Tech*, vol. 11, Oct. 1952, pp. 52–53, 86–88; vol. 11, Nov. 1952, pp. 62–63, 114–116, 118. Introductory survey paper. Pt. I compares various modulation systems. Pt. II describes various methods for measuring frequency.

82. An Accurate Digital Analogue Function Generator, **W. A. Farrand.** *Proceedings*, Electronic Computer Symposium, May 1952, pp. XVI: 1–9. Description and block diagram of a 3-speed servo which obtains sampled digital information for each speed in time sequence. Digital bits are decoded in resistor matrix. Velocity servo provides intermediate output values between samples.

83. Bandwidth and Transmission Performance, **C. B. Feldman, W. R. Bennett.** *Bell System Technical Journal*, vol. 28, July 1949, pp. 490–595.

84. A Magnetic Tape Recording System for Precision Data, **L. L. Fisher.** *Proceedings*, Institute of Radio Engineers, vol. 41, March 1953, p. 424. Abstract of paper describing recording of 0–5,000-cycle range by frequency-modulation technique.

85. A New Coding System for Pulse-Code Modulation, **A. G. Fitzpatrick.** *Proceedings*, Institute of Radio Engineers, vol. 41, March 1953, p. 426. Abstract of paper describing system of sampling, quantizing, and coding by two simple special-purpose tubes and minimum of circuitry. Sampling by beam-deflection tube with ten outputs. Quantizing-coding tube has ten controllable beam positions each with parallel binary-coded output.

86. Electronic Machines for Business Use, **W. B. Floyd.** *Electronics*, May 1950, vol. 23, pp. 66–69. Computers versus business office machines. Data handling and operating requirements.

87. A Transistor Optical Position Encoder and Digit Register, **H. G. Follingstad et al.** *Proceedings*, National Electronics Conference, vol. 8, 1952, pp. 766–75. Description and block diagram of linear shaft position coder using code slide and photo cells actuated by flash tubes. Circuit for pulse regeneration and cyclical-to-straight binary code converter.

88. Data Storage in Three Dimensions, **Jay W. Forrester.** *Whirlwind Memorandum M-70*, Massachusetts Institute of Technology, Cambridge, Mass., April 29, 1947, 6 pp. Short outline of 3-dimensional storage using glow discharge as nonlinear impedance element. Gives broad, preliminary principles.

89. Digital Information Storage in Three Dimensions Using Magnetic Cores, **Jay W. Forrester.** *Journal of Applied Physics*, vol. 22, Jan. 1951, pp. 44–48.

90. Fluorescent and Other Gaseous Discharge Lamps (book), **William E. Forsythe, Elliot Q. Adams.** Murray Hill Books, Inc., New York, N. Y., 1948. Exhaustive treatise on the physics and engineering of fluorescent lamps. First two chapters treat the theory of light and radiation and the discharge through gasses.

91. Measurement of Radiant Energy (book), **William E. Forsythe,** editor. McGraw-Hill Book Company, Inc., New York, N. Y., 1937.

92. Spectral Characteristics of Photographic Light Sources, **William E. Forsythe.** *Journal*, Photographic Society of America, vol. 8, Oct. 1942, pp. 374–84. General discussion of topic with extensive reference to and data of photoflash lamps. Graphs of the effects of applied voltage variation on the light output and color temperatures.

93. The Metrotype System of Digital Recording and Telemetering, **G. E. Foster.** *Electrical Engineering*, vol. 69, May 1950, pp. 427–30. Equipment description for power application. Sampling time 0.8 second. Quantizing by pulse counter and phantastron.

94. The Numeroscope, **Harrison W. Fuller.** *Symposium on Large-Scale Digital Calculating Machinery*, Harvard University Press, Cambridge, Mass., 1948, pp. 238–47. Description of an output device where decimal digits are formed on cathode-ray tube and photographed. Detailed explanation of number formation on cathode-ray tube.

95. Communication Theory and Physics, D.

Gabor. *Philosophical Magazine*, Nov. 1950, vol. 41, pp. 1161–87.

96. L'Enregistrement Magnétique Sonore, **Francois Gallet.** *L'Onde Électrique*, Nov. 1950, vol. 30, pp. 449–57. General paper on magnetic sound recording giving usual theoretical background. Crosstalk from magnetization of two adjacent wires. Coated and impregnated tapes are considered. English summary is in *Science Abstracts*, 1951, vol. 54-B, no. 1520, p. 208.

97. Coincidence Detectors for Binary Pulses, **Clarence Gates.** *Proceedings*, Institute of Radio Engineers, vol. 41, March 1953, p. 426. Abstract only.

98. Manufacturer's Uniform Specification for Special Oscillograph. *Report O.S. 4825, Contract NOrd 10821*, General Electric Company, Schenectady, N. Y., Jan. 22, 1952.

99. PM17D5 Oscillograph Characteristics. General Electric Company, Lynn, Mass.

100. Pulse Width Discriminator, **A. A. Gerlach, D. S. Schover.** *Electronics*, June 1951, vol. 24, pp. 105–07.

101. Problems Involved in Magnetic Tape Recording, **Norman E. Gibbs.** *Proceedings*, Electronic Computer Symposium, May 1952, pp. III: 1–14. Study of causes of pulse dropouts and some corrective measures. Block diagram of test equipment.

102. Zur Lichtemission stromstarker Funkenentladungen, **G. Glaser.** *Optik*, vol. 7, 1950, 33–53, 61–90. Light emission from high current spark discharges.

103. Notes on Digital Coding, **M. J. E. Golay.** *Proceedings*, Institute of Radio Engineers, vol. 37, June 1949, p. 657.

104. Television by PCM, **W. M. Goodall.** *Bell System Technical Journal*, vol. 30, Jan. 1951, pp. 33–49. Describes special features of a pulse code modulation system suitable for television. Signals are sampled from code tube at 10-megacycle rate, transmitting five digits of cyclical code in parallel. Shows pictures of test results with variation of digits and noise.

105. Multi-Stable Magnetic Memory Systems, **John D. Goodell, Tenny Lode.** *Proceedings*, National Electronics Conference, vol. 7, 1951, pp. 378–79. Short note suggesting magnetic intensity modulation recording with counting pulses as output. No details are given.

106. Past, Present and Future Recording Systems, **John D. Goodell.** *Radio and Television News*, vol. 45; *Radio-Electronic Engineering*, vol. 16, May 1951, pp. 11–13A, 27A. Review of concepts that may lead to superior recording systems. Four basic media: physical, chemical, magnetic, and electrostatic. Application to computers. Area, intensity, parallel, and series recording.

107. Design of Apparatus for Speech Stretching, **G. T. Gould.** *Technical Data Digest*, vol. 16, May 1951, pp. 14–19.

108. A Proposal for a Binary Adder Tube Utilizing Beam-Deflection Principles, **D. H. Gridley.** *Report 3802*, Naval Research Laboratory, Washington, D. C., Jan. 25, 1951, 6 pp. Preliminary description of dual-beam tube capable of acting as arithmetic element. One tube required per binary digit.

109. Pulse-Count Modulation, **D. D. Grieg.** *Electrical Communication*, vol. 24, Sept. 1947, pp. 287–96. Almost identical to reference 110.

110. Pulse-Count Modulation System, **D. D Grieg.** *Tele-Tech*, vol. 6, Sept. 1947, pp. 48–52, 98. General description of pulse count modulation transmitting-receiving system. Equations are given for quantization distortion, time sampling distortion, bandwidth signal-to-noise ratio, and threshhold and other design relations. Quantization is accomplished by pulse width modulation, counting of corresponding pulses in straight binary counter. Bibliography.

111. A Binary Quantizer with a Resolution of One Part in a Thousand, **Julian E. Gross.** *M.S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., May 1951, 56 pp. A coarse-fine quantizer. Ten binary digits. One thousand samples per second, 12 clock pulses are counted for five most significant digits. Interval between last clock pulse and coincidence pulse stretched by 32 times and counted again. Detail tests of comparator accurate to 0.1 volt and time amplifier which multiplies by 32. Comparator is based on multiar circuit.

112. Die Wechselbeziehungen zwischen Magnettonband und Ringkopf bei der Wiedergabe, **Walter Guckenburg.** *Funk und Ton*, vol. 4, Jan. 1950, pp. 24–33. Interaction between magnetic tape and reproducing head. Quantitative treatment of effects of self-demagnetization, gap in head, etc. Six references to German literature.

113. Miniaturization of Neon Tubes for Digital Recorders, **Norman R. Gunderson, G. P. Wilson.** *Progress Report 2, TM 687*, United States Naval Ordnance Test Station, Inyokern, Calif., July 1952, 26 pp. Details of experiments and equipment to develop a neon tube for photographic recording. Tube is about 1 inch long and 0.05 inch in diameter. Construction details of various tube types and their difficulties. Time variation of tube characteristics. Tubes were capable of operation at 20 kc with 180-volt peak potential.

114. Miniaturization of Neon Tubes for Digital Recorders, **Norman R. Gunderson.** *Progress Report 3, TM 1262*, United States Naval Ordnance Test Station, Inyokern, Calif., Feb. 24, 1953, 9 pp. Final report of project to develop small neon bulbs with low striking voltage. Construction details of tube with striking potential below 100 volts.

115. A Head for Static Reading of Magnetic Recording, **Donald H. A. Hageman.** *M.S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., June 1951, 102 pp. Description of design of a reading head which measures flux by the change of permeability of the core when flux from the tape passes through it. Theory of static reading. Design and construction details. Test apparatus and procedures. Experimental results. Bibliography.

116. Spin Echoes, **E. L. Hahn.** *Physical Review*, vol. 80, Nov. 15, 1950, pp. 580–94. Description and analysis of nuclear spin systems. Echo patterns might be suitable for storing information very closely.

117. Versatile, Inkless, Multichannel Recorder, **D. W. Halfhill.** *General Electric Review*, vol. 52, Nov. 1949, pp. 39–42.

118. Error Detecting and Error Correcting Codes, **R. W. Hamming.** *Bell System Technical Journal*, vol. 26, April 1950, pp. 147–60. "Systematic" error correcting codes are developed. A theory of error checking and error correcting codes is developed based upon an *n*-dimension unit cube. Formulas and tables are given for selection of optimal codes, which have minimum redundancy for a desired checking and correction feature.

119. The Principles of Optics (book), **Arthur C. Hardy, Fred H. Perrin.** McGraw-Hill Book Company, New York, N. Y., 1st ed., 1932. Standard textbook containing chapter on characteristics of photographic materials.

120. Analysis of Errors in a Phase-Shift Angle Telemetering System, **John V. Harrington.** *CFS Report E 5027*, United States Air Force, Air Materiel Command, Cambridge, Mass., Nov. 1948, 33 pp. Analysis of errors from various sources in a phase-shift modulated system to transmit angle data. Principal error source is nonlinearity of transmission system.

121. A Five-Digit Parallel Coder Tube, **John V. Harrington, K. N. Wulfsberg, G. R. Spencer.** *Proceedings*, Institute of Radio Engineers, vol. 39, March 1951, p. 306. One-paragraph abstract of paper proposing cyclic coding pattern on cathode-ray tube using material of high secondary emission ratio.

122. Proceedings of a Symposium on Large-Scale Digital Calculating Machinery (book), Harvard University Press, Cambridge, Mass., 1948. (Vol. 16 of the Annals of the Computation Laboratory of Harvard University.) Thirty-four papers on computers covering: description of existing machines, computer logic, storage devices, numerical methods, coding, and input-output devices.

123. Bibliography: Counting Circuits and Secondary Emission, **Monica Healea.** *Nucleonics*, vol. 2, March 1948, pp. 66–74. Many titles applicable primarily to atomic physics. No abstracts.

124. Some Distinctive Properties of Magnetic-Recording Media, **R. Herr, B. F. Murphy, W. W. Wetzel.** *Journal*, Society of Motion Picture Engineers, vol. 52, Jan. 1949, pp. 77–88. Experimental results showing the effect of bias adjustment on distortion, frequency response, overload characteristics, and permanency. Curves of spacing loss are given. Tape noise is discussed.

125. Miniature Airborne Telemetering System, **H. M. Hill, Jr.** *Tele-Tech*, vol. 11, Dec. 1952, pp. 68–72, 84, 86, 90, 94, 96. Equipment description and block diagram discussion. Comparison of frequency modulation, amplitude modulation, and pulse modulation and their effects on radio frequency power and bandwidth requirements for given signal-to-noise ratio.

126. Radio Telemetering, **G. L. Hinkley.** *Electronic Engineering*, vol. 21, June 1949, pp. 209–11, 223. Introduction to telemetering covering briefly choice of modulation method, multiplexing, and carrier frequency. Bibliography.

127. Factors Governing the Choice of a Number Base for Use in a Digital Computer, **G. W. Hobbs.** *Report TR 55413*, General Electric Company, Schenectady, N. Y., June 7, 1949, 20 pp. Number theory is used to derive most efficient number bases, mixed and simple, for data writing and for arithmetic operations.

128. Developments in Magnetic Recording, **P. T. Hobson.** *Electronic Engineering*, vol. 19, Dec. 1947, pp. 377–82.

129. Crystal Counters, **Robert Hofstadter.** *Nucleonics*, vol. 4, April 1949, pp. 2–27; vol. 4, May 1949, pp. 29–43. This is not a counting circuit.

130. Crystal Counters, **Robert Hofstadter.** *Proceedings*, Institute of Radio Engineers, vol. 38, July 1950, pp. 726–40.

131. Bibliography on Data Storage and Recording, **Gerhard L. Hollander.** *Technical Memorandum No. 8*, Massachusetts Institute of Technology, Servomechanisms Laboratory, Cambridge, Mass., Feb. 24, 1953. Preliminary, shorter version of this bibliography.

132. Criteria for the Selection of Analogue-to-Digital Converters, **Gerhard L. Hollander.** *Proceedings*, National Electronics Conference, vol. 9, 1953. Classification of analogue-to-digital converters for voltages into ten groups by method of operation. Included are a block diagram for each group and equations relating the approximate space requirements, sampling speed, and precision of each converter.

133. Digital Data Recorder, **Gerhard L. Hollander.** *Technical Memorandum 6897-TM-12*, Massachusetts Institute of Technology, Servomechanisms Laboratory, Cambridge, Mass., 1953, 180 pp. Block diagrams of a recorder capable of storing with 0.5-per-cent precision two channels of 1,000-cycle bandwidth and six channels of 200-cycle bandwidth. Its operating time is 8 minutes; maximum volume specification, 500 cubic inches. The results of a study to store maximum information in minimum space are appended. Detailed treatment of basic concepts of data recording, analogue-to-digital converters, photographic film and magnetic tape storage.

134. Recording and Storing of Information, **Gerhard L. Hollander.** *Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., May 1953, 180 pp. Essentially identical to reference 133.

135. An Evaluation of the Application of New and Old Techniques to the Improvement of Magnetic Recording Systems, **Lynn C. Holmes.** *Proceedings*, National Electronics Conference, vol. 4, 1948, p. 46. One-paragraph abstract of paper presented.

136. Investigation of Magnetic Recording Systems, **Lynn C. Holmes.** *Project RGE-9017*, Stromberg-Carlson Company, Rochester, N. Y., May 2, 1949, 240 pp. Results of study to find magnetic recording methods having higher frequency response, better transient rise time, higher signal-to-noise ratio, lower distortion. Detailed description of all test equipment built—70 pp. Comparison of magnetic recording media and heads. Detailed results of tests to find and evaluate factors limiting frequency response, signal-to-noise ratio, distortion —100 pp.

137. Supersonic Bias for Magnetic Recording, **Lynn C. Holmes, D. L. Clark.** *Electronics*, vol. 18, July 1945, p. 126.

138. A Preliminary Investigation of Transistors for Computer Circuits, **William P. Horton, Robert M. Kalb.** *Contract NObsr-42001*. Engineering Research Associates, Saint Paul, Minn., Oct. 7, 1949, 28 pp. Brief treatment of transistor circuits and their equivalents. Experimental evaluation of a few transistors. Some of the results differ from expected values.

139. Status of High-Speed Digital Computing Systems, **Harry D. Huskey.** *Mechanical Engineer-*

ing, vol. 70, Dec. 1948, pp. 975–78. The various elements of computers including external and internal memories. Speed and storage capacity of all types of memories mentioned. Bibliography.

140. Nonlinear Coil Generators of Short Pulses, **L. W. Hussey**. *Proceedings*, Institute of Radio Engineers, vol. 38, Jan. 1950, pp. 40–44.

141. Semiconductor Diode Gates, **L. W. Hussey**. *Bell System Technical Journal*, vol. 32, Sept. 1953, p. 1137.

142. Novel Recording Accelerometer, **A. S. Iberall**. *Review of Scientific Instruments*, vol. 20, April 1949, pp. 304–07. Small accelerometer of slow chart speed. Spark recording.

143. Coil Galvanometers with Some Novel Features, **G. Ising**. *Arkiv för Matematik, Astronomi och Fysik*, vol. 36-A, paper no. 13, 1948. Abstracted in *Science Abstracts*, vol. 52-B, Feb. 1949, p. 58. High sensitivity galvanometer.

144. Ein lichtelektrisch gesteuerter Fallbügelschreiber, **Lothar Jäger**. *Acta Physica Austriaca*, vol. 4, Dec. 1950, pp. 213–17. A photoelectrically controlled chopper-bar recorder. Describes recording mechanisms for use with reflecting instruments.

145. Barrier-Grid Storage Tube and its Operation, **A. S. Jensen et al.** *RCA Review*, vol. 9, March 1948, pp. 112–33. Exhaustive paper on barrier-grid storage tube.

146. A Phase-Shift Transmission System for Rotational Data, **E. A. Johnson**. *Technical Note 25*, Radar Research Development Establishment, Malvern, England, May 24, 1948, 4 pp.

147. Factors Affecting Spurious Printing in Magnetic Tapes, **S. W. Johnson**. *Journal*, Society of Motion Picture Engineers, vol. 52, June 1949, pp. 619–28. Test results of five types of tape to determine the "echo" effect as a function of time, temperature, and output level of the original recording.

148. Development of Magnet Wire, **William R. Johnson, Max Hansen**. Armour Research Foundation, Illinois Institute of Technology, Chicago, Ill., Feb. 17, 1950, 6 pp. (Final Report). Results of study of magnet-wire materials with the primary aim of developing a conductor with high conductivity and low-temperature coefficient of resistance. Results are negative.

149. Spectral Distribution of Sensitivity of Photographic Material, **L. A. Jones, Otto Sandrik**. *Journal of the Optical Society of America and Review of Scientific Instruments*, vol. 12, April 1926, pp. 401–16.

150. The "Filmgraph" Sound Recording System, **J. H. Jupe**. *Electronic Engineering*, vol. 19, Dec. 1947, p. 389. One-page description of recorder making nonvolatile, eraseable recordings on film by indenting with a stylus. Lateral track width less than 0.003 inch. Response flat from 75–5,000 cycles. Film speed 5 inches per second and 100 tracks are possible on a single film. Unit capable of recording 300 hours, weighs 30 pounds. Operating cost considerably below steel or wax media.

151. Recommended Programs for Improvement and Standardization of Computer Tubes, **Joseph Kelar**. *Report PX 29483, Contract NObsr-42001*, Engineering Research Associates, Saint Paul, Minn., March 19, 1951, 6 pp. Short paper on need for better tubes.

152. Data Recording on Magnetic Tape, **Leo G. Killian**. *Electronic Industries and Electronic Instrumentation*, vol. 2, April 1948, pp. 3–5, 31. System description and block diagrams of air-borne magnetic recorder capable of withstanding 75 g. Data is recorded by frequency modulation.

153. Static Magnetic Memory for Low-Cost Computers, **M. Kincaid, J. M. Alden, R. B. Hanna**. *Electronics*, vol. 24, Jan. 1951, pp. 108–11.

154. A Ratio Quantizer, **A. M. King, J. T. McLane, G. G. Myers**. *Report 4085*, Naval Research Laboratory, Washington, D. C., Jan. 2, 1953, 59 pp. Theory and block diagrams of dividing circuit. Division is performed by charging a central condenser by quanta of charges proportional to the two unknown voltages. A switching network selects charges from one or the other unknown to keep the voltage on the central condenser zero. The number of quanta from each unknown is counted.

155. Repeller Storage Tube, **H. Klemperer, J. T. DeBettencourt**. *Electronics*, vol. 21, Aug. 1948, pp. 104–06.

156. Pulse Generation and Shaping at Microwave Frequencies, **W. A. Klute**. *Bell Laboratories Record*, vol. 29, May 1951, pp. 216–20.

157. Storage Tubes and Their Basic Principles (book), **M. Knoll, B. Kazan**. John Wiley and Sons, New York, N. Y., 1952.

158. Preparation and Characteristics of Magnetic Recording Surfaces, **Edward D. Korhone et al.** *Contract NObsr-42001*, Engineering Research Associates, Saint Paul, Minn., Jan. 7, 1949, 19 pp. Description and evaluation of surfaces prepared by spraying. Description and block diagram of test methods. For pulse densities between 50 and 100 pulses per inch, curves are given for resolution, signal strength at pickup head, and signal-to-surface-noise ratio.

159. Frequency Response of Magnetic Recording, **Otto Kornei**. *Electronics*, vol. 20, Aug. 1947, pp. 124–28. Discussion of magnetic properties of recording medium such as demagnetization, penetration, gap effect, and tape velocity. Description of electroplated wire and powder-coated tape.

160. Investigations of New Magnetic Recording Media, **Otto Kornei et al.** *OSRD Report 5325*, Brush Development Company, Cleveland, Ohio, June 30, 1945, 70 pp. Detailed report on development of new tapes. Details of materials and plating methods. Test equipment.

161. Survey of Magnetic Recording, **Otto Kornei**. *Symposium on Large-Scale Digital Calculating Machinery*, Harvard University Press, Cambridge, Mass., 1948, pp. 223–37. Good qualitative summary of magnetic recording phenomena. Frequency response curves of magnetic wire are analyzed graphically. Comparison of solid and powdered tapes. Bibliography.

162. Trends and Developments in Magnetic Recording Heads, **Otto Kornei**. *Brush Strokes*, Brush Development Company, Cleveland, Ohio, March 1952, pp. 1–7. Detailed design description of Brush types *BK-1500* and *BK-1300* recording heads.

163. Ein Gesetz fur die Schwärzungskurven bei Mischfarben, **A. van Kreveld**. *Zeitschrift für wissenschaftliche Photographie, Photophysik und Photochemie*, vol. 32, Jan. 1934, pp. 222–30. Describes a law concerning density curves due to mixed colors. Theoretical derivation and experimental verification that the densities due to two different light frequencies are additive at simultaneous exposure. Excellent paper.

164. Das Photographische Summengesetz und sein Gültigkeitsbereich, **A. van Kreveld**. *Physica*, vol. 1, 1934, pp. 60–77. The summation law and its region of validity. Summary of the summation law. New generalization from additional experimental data particularly as applied to the Schwarzschild-exponent and Soldberg's statistical law of grain sensitivity.

165. Kanalkapazität und Laufzeit, **Karl Küpfmüller**. *Archiv der Elektrischen Übertragung*, vol. 6, July 1952, pp. 265–68. Shannon's formula is approximately true only for transmission times greater than 20/bandwidth. A formula for channel capacity for finite transmission time is derived. Numerical results are presented in graphical and tabular form.

166. Cathode-Follower Gate Circuit, **Jerome Kurshan**. *Review of Scientific Instruments*, vol. 18, Sept. 1947, pp. 647–49. Circuits, design equations and procedure for a gate circuit using a dual-triode with common cathode load resistor.

167. Limitations of Magnetic Tape, **W. S. Latham**. *Audio Engineering*, vol. 36, Sept. 1952, pp. 19–20, 68–69. The effect of tape imperfections and losses as a function of number of playbacks, up to 2,000. Microphotographs, oscillographs, and curves show enlarging of minor imperfection and effect of number of playbacks on frequency content of a given tape. High frequencies, above 100 cycles, attenuate with replays, while below 50 cycles they are actually boosted.

168. A High Speed Electronic Curve Printer, **John George Lawton**. *M.S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1951, 106 pp.

169. Ultra Speed Recording for Acoustical Measurements, **C. J. LeBel, J. Y. Dunbar**. *Journal*, Acoustical Society of America, vol. 23, Sept. 1951, pp. 559–63. Methods of recording reverberation decay characteristics by new ultraspeed technique; higher pen speeds.

170. Noise Effects in FM-FM Telemetering, **Frank W. Lehan**. *CIT/JPL Memorandum 4-47* California Institute of Technology, Pasadena Calif., Feb. 14, 1949, 11 pp. Five-page text. Mathematical manipulation of noise and application to California Institute of Technology telemetering system.

171. Transistor Oscillator for Telemetering' **Frank W. Lehan**. *Electronics*, vol. 22, Aug. 1949' pp. 90–91.

172. A Multi-Channel Oscillograph Utilizing Color Film, **W. J. Leiss et al.** *Instruments*, vol. 20, Aug. 1947, pp. 709–11. Equipment description of a 6-channel oscillograph using mirror galvanometers with color filters and color film to separate the traces.

173. High-Speed Information-Storage Devices and Their Applications in Communication Engineering, **A. J. Lephakis**. Massachusetts Institute of Technology, Cambridge, Mass., Sept. 1948, 134 pp. (Electrical Engineering Department Seminar Series.) Detailed description of many flip-flops and counters, transmission line and cathode-ray tube storage. Extensive bibliography.

174. Storage of Pulse Coded Information, **A. J. Lephakis**. *M.S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1949, 142 pp. Description and comparison of various storage media as flip-flops, ultrasonic delay lines, storage tubes, and magnetic tapes and drums. Detailed description of final storage system using Massachusetts Institute of Technology tubes.

175. A High-Precision Analog-to-Digital Converter, **Bernard Lippel**. *Proceedings*, National Electronics Conference, vol. 7, 1951, pp. 206–15. Description and some theoretical material on coding wheel using photocell pickup and cyclical code. Fifteen-digit wheel. Digital servo is mentioned.

176. Interconversion of Analog and Digital Data in Systems for Measurement and Control, **Bernard Lippel**. *Proceedings*, National Electronics Conference, vol. 8, 1952, pp. 636–46. Survey of analogue-to-digital and digital-to-analogue converters. Classification of coders and decoders as counting, reading, and weighing types.

177. A Systematic Survey of Coders and Decoders, **Bernard Lippel**. *Convention Record*, Institute of Radio Engineers, pt. 8, 1953, pp. 109–19. Paper classifying coders and decoders into three basic groups.

178. A Systematic Survey of Coders and Decoders, **Bernard Lippel**. *Proceedings*, Institute of Radio Engineers, vol. 41, March 1953, p. 425. Abstract of reference 177.

179. Exposé Simplifié de la Théorie Informationelle de Shannon en Vue de son Application aux Problèmes de Télécommande, Télémesure ou Mesure, **J. Loeb**. *Annales des Télécommunications*, vol. 6, March 1951, pp. 67–76. Coding of messages for equal and unequal probabilities. Effects of coding.

180. Une Théorie 'Informationnelle' de la Mesure et de la Télémesure, **J. Loeb**. *Annales des Télécommunications*, vol. 6, April 1951, pp. 90–97. Application of Shannon and Hartley formulas to electric measurement of voltages. Precision requirements.

181. High-Speed Direct-Trace Recording Methods, **Roger Brown Loucks**. *Instruments*, vol. 24, Jan. 1951, pp. 30–34. Recording physiological actions as eye-blinks by: 1. potentiometer slider operated by eye-lid, 2. light shutter controlling photocell. Gives circuit of fixed stylus recorders using Teledeltos paper. Balance of paper suggests various input elements such as potentiometers, mirror galvanometer and photocells, cathode-ray commutator tube.

182. Magnetische Schallaufzeichnung mit Filmen und Ringköpfen, **Heinz Lübeck**. *Akustische Zeitschrift*, vol. 2, Nov. 1937, pp. 273–95. A fundamental investigation of various factors in magnetic tape recording. Excellent analytical approach and experimental verifications.

183. The Construction and Characteristics of a Non-Contact Electromagnetic Recording and Reproducing Head, **G. Ward Lund**. *Contract NObsr-42001*, Engineering Research Associates, Saint Paul, Minn., Aug. 7, 1950, 27 pp. Description and photographs of a head suitable for production. Written from point of view of fabrication.

184. The Dual Element Magnetic Head, **G. Ward Lund**. *Report PX29697, Contract NObsr-42001*, Engineering Research Associates, Saint Paul, Minn., Jan. 16, 1952, 35 pp. Description,

detailed design, and experimental evaluation of a head containing a read and a write element.

185. Report on Telemetering in Aero Research, **M. McLarren.** *Aviation Week*, vol. 51, Nov. 28, 1949, pp. 24, 28, 31–32.

186. Photographic Aspects of the Theory of Three-Color Reproduction, **David L. MacAdam.** *Journal*, Optical Society of America, vol. 28, Nov. 1938, pp. 399–418. Contains, among other material, a short review of van Kreveld's law and the concept of spectral sensitivity.

187. On the Energy-Spectrum of an Almost Periodic Succession of Pulses, **G. G. Macfarlane.** *Proceedings*, Institute of Radio Engineers, vol. 37, Oct. 1949, pp. 1139–47. Discussion: vol. 38, Oct. 1950, pp. 1212–13. Derivations of power-density spectra for equally spaced pulses of random amplitude, for constant amplitude pulses spaced almost equally, and for regular pulses of unequal length.

188. Multichannel Analog Input-Output Conversion System for Digital Computer, **M. L. MacKnight, P. A. Adamson.** *Convention Record*, Institute of Radio Engineers, pt. 7, 1953, pp. 2–6. Description and block diagrams of a miniaturized coding and decoding system. Asynchronous operation from computer by a storage drum memory. Ten coding and four decoding channels are sampled ten times per second each. One coding channel used for calibrating analogue-to-digital converter. Precision, nine digits. Sweep-timing coder and multiar comparator are used.

189. Multichannel Magnetic Recording Device, **Josiah Macy, Jr.** Massachusetts Institute of Technology, Group Networks Laboratory, Cambridge, Mass., 1951, 6 pp. Equipment description capable of recording 600 samples per second of on-off information.

190. Ein neuer electrischer Schnellschreiber und seine Anwendung in der elektromedizinischen Messtechnik, **K. Mall, O. Grimm.** *VDE-Fachberichte*, vol. 14, 1950, pp. 156–60. A new electric recorder and its application in medical measurement. Description of an oscillograph using heat-sensitive paper. Frequency response 0.1 to 120 cycles. Using amplifiers included 1 millivolt produces 15-millimeter deflection.

191. Interference Effects in Magnetic Recording Heads, **Arthur H. Mankin.** *Proceedings*, National Electronics Conference, vol. 8, 1952, pp. 108–12.

192. Compression de Fréquences, **P. Marcou.** *Annales des Télécommunications*, vol. 5, Oct. 1950, pp. 321–37. Mathematical analysis of problems connected with signal compression by sampling. Conditions for satisfactory compression and transmission on narrow channels.

193. Gas-Discharge Gaps for Data Storage in Electronic Computers, **R. F. Markel.** *Whirlwind Report R-140*, Massachusetts Institute of Technology, Cambridge, Mass., July 1948.

194. Preliminary Classified Bibliography on Random Disturbances. *DACL Memorandum M-6.16-2*, Massachusetts Institute of Technology, Cambridge, Mass., Dynamic Analysis and Control Laboratory, Sept. 14, 1948, 12 pp. Approximately 100 references without abstracts by subject headings. 1. Descriptive and quantitative information on random disturbances; 2. system studies including the effect of random disturbances; 3. system design for optimum performance.

195. An FM/FM Telemetering System, **W. J. Mayo-Wells.** *Instruments*, vol. 23, July 1950, pp. 717–18. Block diagram and description of specific system.

196. An Experimental Multichannel Pulse Code Modulation System of Toll Quality, **L. A. Meacham, E. Peterson.** *Bell System Technical Journal*, vol. 27, Jan. 1948, pp. 1–43.

197. Photography (book), **C. E. Kenneth Mees.** The Macmillan Company, New York, N. Y., 2nd ed., rev., 1951. Fundamentals of photography written in a popular style.

198. The Theory of the Photographic Process (book), **C. E. Kenneth Mees.** The Macmillan Company, New York, N. Y., 1942. Advanced treatment of photography. Summary of results with many references to the original papers. Excellent bibliography.

199. Long Range Multi-Channel Telemetering System, **J. Mengel.** *Instruments*, vol. 23, Jan. 1950, pp. 70–72. Pulse-time modulation of 30 channels with 312.5 samples per second per channel. Equipment description with block diagrams. Direct voltage levels are transmitted by position of single pulse in a 100-microsecond interval.

200. Fundamentals of the Automatic Telephone Message Accounting System, **John Meszar.** *AIEE Transactions*, vol. 69, 1950, pp. 255–69. Bell system automatic subscriber charge recording system charges all calls and processes data in accounting. Economic background, coding technique, features of recording and processing machinery.

201. Bibliography of Magnetic Amplifier Devices and the Saturable Reactor Art, **James G. Miles.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 2104–23. Nine hundred references to papers and significant patents. Patents are well indexed.

202. Bibliography: Magnetic Amplifier Devices and the Saturable Reactor Art, **James G. Miles.** *Contract NObsr-42001*, Engineering Research Associates, Saint Paul, Minn., Sept. 1, 1950, 1st ed., 64 pp. Superseded by later edition.

203. Bibliography: Magnetic Amplifeir Devices and the Saturable Reactor Art, **James G. Miles.** *Report PX29520*, Engineering Research Associates, Saint Paul, Minn., Aug. 31, 1951, 2nd ed., 71 pp. Nine hundred forty-five references without abstracts. Essentially identical to reference 201.

204. Saturable-Core Reactors as Digital Computer Elements, **James G. Miles.** *Contract NObsr-42001*, Engineering Research Associates, Saint Paul, Minn., June 17, 1949, 63 pp. Background on saturable cores and their application to single-core flip-flops, high-speed 2-core flip-flops, counters, and gates. Theoretical and quantitative information. Bibliography of 93 titles on magnetic amplifiers.

205. Saturable Reactors as Substitutes for Electron Tubes in High Speed Digital Computers, **James G. Miles.** Engineering Research Associates, Saint Paul, Minn., March 21, 1951, 20 pp. Short paper describing use of small magnetic cores for flip-flops, counters, and gates. Experimental results show flip-flops operating at repetition rates of 400,000 pulses per second.

206. The Mechanism of the Supersonic Bias, **Angelo Montani.** *Electrical Engineering*, vol. 68, June 1949, p. 511. One-page digest using push-pull class-B amplifier analogy to explain effect of bias.

207. Magnetic and Phosphor Coated Discs, **Benjamin L. Moore.** *Symposium on Large-Scale Digital Calculating Machinery*, Harvard University Press, Cambridge, Mass., 1948, pp. 130–32. Brief description of proposed storage methods. Disc coated with phosphorescent material would rotate with write and read devices spaced 180 degrees apart. Magnetic disc or drum can store information statically.

208. High-Speed Photography, **Kenneth Morgan.** *Interchemical Review*, vol. 11, 1952, pp. 59–70. Survey of high-speed cameras. Description of equipment and techniques. Typical results.

209. Stabilized Time-Division Multiplier, **C. D. Morrill, R. V. Baum.** *Electronics*, vol. 25, Dec. 1952, pp. 139–41.

210. Methods of Calibrating Frequency Records, **R. C. Moyer et al.** *Proceedings*, Institute of Radio Engineers, vol. 38, Nov. 1950, pp. 1306–13.

211. Recording Demagnetization in Magnetic Tape Recording, **O. William Muckenhirn.** *Proceedings*, Institute of Radio Engineers, vol. 39, Aug. 1951, pp. 891–97. Theoretical analysis and experimental work proposing the recording demagnetization theory to explain certain performance characteristics of magnetic tapes.

212. Design Features of the Era 1101 Computer, **F. C. Mullaney.** *Electrical Engineering*, vol. 71, Nov. 1952, pp. 1015–18. Brief general computer description. Magnetic drum memory is used.

213. Flutter Compensation for FM-FM Telemetering Recorder, **J. T. Mullin.** *Proceedings*, Institute of Radio Engineers, vol. 41, March 1953, p. 424. Abstract of paper recommending addition of high-frequency tone to each channel during recording. All channels are modulated by a high-frequency carrier. By proper demodulation signal is restored.

214. Radio Telemetry, **M. N. Nichols, L. L. Rauch.** *Review of Scientific Instruments*, vol. 22, Jan. 1951, pp. 1–29. Detailed, mainly theoretical discussion of factors in choosing modulation system in aircraft multichannel telemetering. Forms of multiplexing. Fluctuation and impulse noise, nonlinearity, frequency response, crosstalk relations. Instrumentation and recording technique. Bibliography. Has section on air-borne recording and efficient coding.

215. An Analog to Digital Converter Using Double Comparison, **C. F. O'Donnell,** *M.S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1951. Design of circuit using Miller integrators to start and stop counting clock pulses. Accuracy 0.002 per unit, speed 100 samples per second, range ±100 volts. Paper treats primarily design of sweep circuit and comparators.

216. The Philosophy of PCM, **B. M. Oliver, J. R. Pierce, C. E. Shannon.** *Proceedings*, Institute of Radio Engineers, vol. 36, Nov. 1948, pp. 1324–31.

217. Photographic Methods for Handling Input and Output Data, **R. D. O'Neal.** *Symposium on Large-Scale Digital Calculating Machinery*, Harvard University Press, Cambridge, Mass., 1948, pp. 260–66. Preliminary description of Eastman Kodak input-output reader.

218. Das Metallpapier-Registrierverfahren, **Alfred Ortlieb.** *Elektrotechnische Zeitschrift*, vol. 71, Dec. 1, 1950, pp. 653–56. The metal-paper recording method. Comprehensive paper giving design and experimental curves on an oscillograph using paper with a thin metal coating which is burned away by electric current. Process features low coefficient of friction and good resolution. Pen has low inertia. Parallel, fixed electrodes can write lines 1.5 millimeters apart.

219. Pulse Code Modulation Systems, **A. J. Oxford.** *Proceedings*, Institution of Radio Engineers, vol. 13, July 1952, pp. 281–86. Also *Proceedings*, Institute of Radio Engineers, vol. 41, July 1953, pp. 859–65. Description of two novel coders and decoders. Quantizing and digital-to-analogue conversion by addition and subtraction of charge on a condenser. Good reliability and precision with minimum equipment is claimed. Brief review of pulse code modulation principles is included.

220. Dynamic Binary Counter with Analog Read-Out, **Leroy Packer.** *Proceedings*, Institute of Radio Engineers, vol. 41, March 1953, p. 405. Abstract only.

221. Germanium Photodiodes Read Computer Tapes, **Leroy Packer, William J. Wray, Jr.** *Electronics*, vol. 25, Nov. 1952, pp. 150–51.

222. Digital Computer Switching Circuits, **C. H. Page.** *Electronics*, Sept. 1948, vol. 21, pp. 110–18. Elementary description of the operation of digital computers with emphasis on the switching operations.

223. Information Theory, **C. H. Page.** *Journal* Washington Academy of Sciences, vol. 41, Aug. 1951, pp. 245–49.

224. Comparison of Modulation Methods, **R. M. Page.** *Convention Record*, Institute of Radio Engineers, pt. 8, 1953, pp. 15–25.

225. Specialized Instrumentation for Flight Testing, **Joseph P. Paine.** *Instruments*, vol. 20, Jan. 1947, pp. 30–34. Primarily concerned with aircraft instrumentation. Recording by photographs of instrument dials.

226. Registrierinstrumente (book), **A. Palm.** Springer Verlag, Berlin, Germany, 1950.

227. A Coincident-Current Magnetic Memory Unit, **W. N. Papian.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1950, 83 pp.

228. A Direct-Writing Oscillograph Utilizing Servomechanism Techniques, **R. Perley.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1951.

229. Test Generator for Pulse Coded System, **D. H. Pickens, A. A. Gerlach.** *Radio and Television News*, vol. 46, July 1951, pp. 14–15.

230. Communication Theory, **J. R. Pierce.** *Physics Today*, vol. 4, May 1951, pp. 6–12.

231. High-Speed Direct-Inking Recording System, **M. A. Pomerantz, R. C. Pfeiffer.** *Journal*, Franklin Institute, vol. 248, Oct. 1949, pp. 305–10. Modification of brush recorder to operate at high paper speed (9 centimeters per second) for extended time, $2^1/_2$ hours. Paper capacity has been increased and mechanical changes produce higher speed.

232. An Electro-Mechanical Reader for Oscillograph Data, **John L. Preston.** *Technical*

Memorandum No. 9, Contract NOrd-11799, Servomechanisms Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., June 1, 1953, 18 pp. Description and circuits of a data reader which reduced oscillograph records with 1-per-cent precision and compensated for variations in excitation voltages.

233. Electrostatically Induced Permanent Memory, **Charles F. Pulvary.** *Journal of Applied Physics*, vol. 22, Aug. 1951, pp. 1039–44. Ferroelectric substance (barium titanate) for storing information can be used in audio-frequency range or higher.

234. The Notched-Disk Memory, **J. Rabinow.** *Electrical Engineering*, vol. 71, Aug. 1952, pp. 745–49.

235. Methods of Modulation and Multiplexing. Radio Corporation of America, RCA Victor Division, Camden, N. J., June 20, 1949, 87 pp. (Final Report, *Contract W-36-039-sc-32075.*) Final report of 3-year theoretical and experimental study. Contains detailed proposed system description (96-channel) and theoretical developments. Refers to earlier quarterly reports for theoretical studies not included.

236. The Selectron—A Tube for Electrostatic Storage, **J. A. Rajchman.** *Mathematical Tables and Other Aids to Computation*, vol. 2, Oct. 1947, pp. 359–61. Short communication describing the selectron.

237. Transistors and Transistor Circuits, **Herbert J. Reich.** *Electrical Manufacturing*, vol. 50, Nov. 1952, pp. 106–12, 324–28.

238. Error Probabilities of Binary Data Transmission Systems in the Presence of Random Noise, **Siegfried Reiger.** *Convention Record*, Institute of Radio Engineers, pt. 8, 1953, pp. 72–79.

239. A Magnetic Record-Reproduce Head, **M. Rettinger.** *Journal*, Society of Motion Picture and Television Engineers, vol. 55, Oct. 1950, pp. 377–90. Discusses constructional principles of *MI-10794* head. Includes equations and graphs showing effects of gap size, lamination thickness, etc., on head inductance, sensitivity, and spacing loss.

240. High-Speed Pulse Recording on Magnetic Tape, **E. S. Rich.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1948, 120 pp. Also *Whirlwind Report R-159*, Massachusetts Institute of Technology, Cambridge, Mass., April 6, 1949. Tests and required apparatus to determine pulse characteristics of tape. Sections I–III give background to pulse recording. Section IV describes construction of equipment. Section V analyzes the data in many forms. Good sets of curves, many nondimensionalized. Bibliography.

241. Magnetic Recording—Its Use for Storage of Information in Electronic Computers, **E. S. Rich.** *Whirlwind Report R-124*, Massachusetts Institute of Technology, Cambridge, Mass., Sept. 17, 1947, 26 pp. Discussion of magnetic recording. Circuits for a *B-H* tracer. Test results obtained on different tapes running at eight inches per second.

242. NACA VGH Recorder, **N. R. Richardson.** *NACA Technical Note 2265*, National Advisory Committee for Aeronautics, Feb. 1951, 16 pp. Also *Aviation Week*, vol. 54, June 18, 1951, pp. 39, 40, 44, 47. Compact flight instrument for time-history records of three quantities for continuous periods of 100 hours. Records are on photographic paper.

243. The Index Marking of Magnetic Tape, **Theodore L. Roess.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., May 16, 1952, 82 pp. Description and circuit diagrams of equipment capable of detecting and marking rapidly tape defects. Result of tests, using this equipment, on frequency of dropouts of various magnitudes. Effect of dust on tape.

244. Static Reading of Magnetically Stored Digital Information, **S. M. Rubens.** *Report PX-29501, Contract NObsr-42001*, Engineering Research Associates, Saint Paul, Minn., Aug. 21, 1951, 25 pp. Construction and description of circuits of a static reading head with a saturable core section. Test results using track width of 0.080 inch and cell densities of 50 to 200 bits per inch for return-to-zero recordings. Interaction between adjacent channels is small, but requires further study.

245. A Method of Pulse-Code Modulation, **Roy Pines Sallen.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1949, 120 pp. Development of two pulse-code modulation systems. Design considerations and circuit description of various components. The binary system, straight and cyclical, and the mathematical proof of the conversions. Consideration of various coding systems. Coder produces cyclical code by successive rectification. Code is available instantaneously, but coder has low sampling rate.

246. Study of Pulse Code Modulation, **F. M. Sanger, Jr.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., Sept. 1948. Circuit diagrams for a pulse-code-modulation decoder. Some experimental results in presence of noise.

247. Spectrograph: Construction, Operation, and Results, **Raymond C. Sangster.** *A Study of Organic Scintillators. Ph.D. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1951, pp. 110–30. Description of a simple spectrograph used to obtain automatically recorded spectra of fluorescent sources. Accuracy ±5 angstroms.

248. An Electronic Scaler for Telemetered Data Reduction, **William F. Santelmann, Jr.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1951, 104 pp. Device quantizes direct voltages and prints resulting 3-digit number on paper strip at up to 100 values per second. Scaling by phantastron width modulation gating of pulses which are counted in biquinary ring counters. Diode matrix controls printer which consists of three wires printing five times per character.

249. Conversion of Binary Pulse Code to Voltage Amplitude, **E. W. Sard.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1948.

250. A Necessary and Sufficient Condition for Unique Decomposition of Coded Messages, **A. A. Sardinas, G. W. Patterson.** *Proceedings*, Institute of Radio Engineers, vol. 41, March 1953, p. 425. Abstract only.

251. Survey of Tape Drive Systems, **H. H. Sarkissan.** *Proceedings*, Electronic Computer Symposium, May 1952, pp. IV: 1–5. Brief description of desirable characteristics of a tape drive.

252. An Analog-to-Digital Converter, **A. D. Scarbrough.** *Report M-58*, Hughes Aircraft Company, Culver City, Calif., Aug. 3, 1953, 9 pp. Brief description of a shaft position coder using cams to develop binary code. Transition errors are minimized by using two switches per cam.

253. Xerography: A New Principle of Photography and Graphic Reproduction, **R. M. Schaffert, C. B. Oughton.** *Journal*, Optical Society of America, vol. 38, Dec. 1948, pp. 991.

254. A Direct-Writing Feedback Recorder, **Ralph Scheidenhelm.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1951, 80 pp. Detailed design information for a pen recorder using a variable capacitor as feedback position element. Static error 1 per cent of full scale. Maximum dynamic error in response to 10-cycle and 20-cycle sinusoidal signals is 5.6 per cent and 15 per cent respectively.

255. Cathode Ray Recorders for Missile Application, **C. H. Schlesman.** *Photographic Engineering*, vol. 3, 1952, pp. 78–88. Describes, diagrams, and illustrates multichannel recorders for flight-monitoring. Photographic recording from a 3-cathode-ray tube provides extreme flexibility and permits channel multiplexing with high-speed electronic switching.

256. A Magnetic Tape Oscillograph for Power System Analysis, **E. C. Schurch, F. R. Schleif.** *Electrical Engineering*, vol. 70, Nov. 1951, pp. 993–97. Oscillograph records conditions on power line continuously; normal conditions can be erased. Modulation by frequency modulation.

257. Beam Deflection Tube for Coding in Pulse Code Modulation, **R. W. Sears.** *Bell Laboratories Record*, vol. 26, Oct. 1948, pp. 411–15.

258. Electron Beam Deflection Tube for Pulse Code Modulation, **R. W. Sears.** *Bell System Technical Journal*, vol. 27, Jan. 1948, pp. 44–57. Paper on quantizer operating on cathode-ray tube principle. Quantizes seven digits in space domain at high speed.

259. Synchronous Recording on 1/4-Inch Magnetic Tape, **Walter T. Selsted.** *Journal*, Society of Motion Picture and Television Engineers, vol. 55, Sept. 1950, pp. 279–84. Description of equipment to synchronize movie film with a sound track on standard 1/4-inch magnetic tape.

260. An Improved Technique for High-Speed Photography, **D. A. Senior, C. O. J. Grove-Palmer.** *British Journal of Applied Physics*, vol. 3, Oct. 1952, pp. 318–21. Discussion of high-speed motion picture technique. Photographs are taken at 4,000 frames per second. Primarily description of control and timing circuits to properly synchronize Fastax camera. Underwater photography with exposure times to 80 microseconds. Gives formula for light required as function of lens aperture, light path, exposure time, and emulsion speed.

261. Communication in Presence of Noise, **Claude E. Shannon.** *Proceedings*, Institute of Radio Engineers, vol. 37, Jan. 1949, pp. 10–21.

262. A Mathematical Theory of Communication, **Claude E. Shannon.** *Bell System Technical Journal*, vol. 27, July 1948, pp. 379–423, Oct. 1948, pp. 623–57.

263. Mathematical Theory of Communication (book), **Claude E. Shannon, W. Weaver.** University of Illinois Press, Urbana, Ill., 1949.

264. Relations Between the Sensitometric and the Size-Frequency Characteristics of Photographic Emulsions, **Ludwik Silberstein, A. P. H. Trivelli.** *Journal*, Optical Society of America, vol. 28, Nov. 1938, pp. 441–59. Theoretical and empirical considerations show that the speed of an emulsion is approximately proportional to the average grain size.

265. Ink Recorders, **G. Simon.** *Siemens Zeitschrift*, vol. 25, July 1951, pp. 141–49. Construction features and performance of Siemens pen recorders.

266. A Digital Electronic Correlator, **Henry E. Singleton.** *Proceedings*, Institute of Radio Engineers, vol. 38, Dec. 1950, pp. 1422–28. Block and circuit diagrams of the correlator and some of its important components. Contains circuits of a sweep-type analogue-to-digital converter with a 10-digit precision and 200-microsecond sampling rate. Binary counter operated reliably at 10 megacycles.

267. The Sadic, A Precision Analog-Digital Converter, **Robert L. Sink, Geo. M. Slocomb.** Consolidated Engineering Corporation, Pasadena, Calif., 6 pp. Description and block diagrams of electromechanical analogue-digital converter with ±0.1-per-cent accuracy and approximately 1/2-cycle frequency response at maximum amplitude. One-microvolt increments can be sensed by using calibrated amplifier before quantizing. Quantizing by differential-counting feedback-subtraction using stepping switches and precision resistors in Thompson-Varley bridge circuit.

268. Devices for Conversion Between Analog Quantities and Binary Pulse-Coded Numbers, **R. L. Sisson, A. K. Susskind.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., Jan. 1950, 130 pp. Description, circuit diagrams, and tests on a digital coder and a decoder. In the coder a synchro produces phase-modulated signal. The phase shift is detected by counting clock pulses in a binary counter.

269. An Analog-to-Digital Converter with an Improved Linear-Sweep Generator, **Dean W. Slaughter.** *Convention Record*, Institute of Radio Engineers, pt. 7, 1953, pp. 7–12. Block diagrams of a sweep-timing coder with a claimed precision of 12 binary digits and pulse intervals of 3 microseconds. Operational amplifiers are used in sweep circuit and comparator. Equations for these two circuits are derived.

270. An Analog to Digital Converter with an Improved Linear Sweep Generator, **Dean W. Slaughter.** *Proceedings*, Institute of Radio Engineers, vol. 41, March 1953, p. 405. Abstract of reference 269.

271. Coding by Feedback Methods, **B. D. Smith.** *Proceedings*, Institute of Radio Engineers, vol. 41, Aug. 1953, pp. 1053–58. Essentially a condensed version of reference 272.

272. Pulse-Code Modulation Method, **Blanchard D. Smith, Jr.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1948, 61 pp. Discussion of pulse code modulation. Design of coder using flip-flops and subtraction principle with single discriminator. Mathematical investigation of quantizing error. Differential-counting method and sequential-error method mentioned. Sampling rate 10 kc which could be extended to 30 kc. Five binary digits per sample.

273. Pulse-Code Modulation Method, **Blanchard D. Smith, Jr.** *Report 23*, Massachusetts In-

stitute of Technology, Cambridge, Mass., Dec. 1948, 29 pp. Condensation of reference 272.

274. Train Performance Calculator, **S. V. Smith.** *Railway Age*, vol. 130, March 26, 1951, pp. 32–33.

275. A Pulse-Code Modulation System, **W. R. Smith-Vaniz, Jr.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., June 1952, 51 pp. Theory and circuits for pulse-code modulation system employing companding. Companding is accomplished in successive-approximation coder. Signal-to-noise ratio 30 decibels for a dynamic range of 20 decibels.

276. A New Method of Recording Flight Data, **Henry Sostman.** *Instruments*, vol. 23, Jan. 1950, pp. 36–37. Description of air-borne, multiple-channel direct-writing oscillograph using thermosensitive paper. D'Arsonval galvanometer movement. Data tabulated.

277. Bibliography on Telemetering, 1924–1947. Special Libraries Association, Cleveland, Ohio, 4 pp. Thirty-three references without abstracts from widely available publications. Apparently all refer to power telemetering.

278. Magnetic Recording Tape. I-II, **H. G. M. Spratt.** *Wireless World*, vol. 57, 1951, pp. 88–91, 149–51. Basic characteristics for specific tape speed and effect of finite gap length.

279. Conversion of Shaft Position to Binary Code, **H. P. Stabler.** *Whirlwind Report R-129*, Massachusetts Institute of Technology, Cambridge, Mass., Oct. 15, 1947, 26 pp. Comparison of various methods to convert shaft position to binary code accurate to one in 100 to 500 and 1,000 to 5,000 parts. Mechanical switches on the function level, and count-down methods using variable frequency oscillator or sweep timing are compared on linearity, stability, and speed. Digital-to-analogue conversion is briefly considered.

280. Reversible Binary Counter and Shaft Position Indicator, **H. P. Stabler.** *Technical Report 3*, Massachusetts Institute of Technology, Research Laboratory of Electronics, Cambridge, Mass., March 3, 1947, 8 pp. A binary counter responds to "add" and "subtract" pulses in unit steps. Resolving time 10^{-6} second and better. Shaft position indicator employs disk which energizes "add" or "subtract" photocell depending on direction of rotation.

281. Design Principles of Amplitude-Modulated Subcarrier Telemeter System, **C. K. Stedman.** *Proceedings*, Institute of Radio Engineers, vol. 36, Jan. 1948, pp. 36–41.

282. Oscillograph Trace Identification, **R. Stott.** *CIT/JPL Memorandum 4-35*, California Institute of Technology, Pasadena, Calif., July 27, 1948, 3 pp. Brief description and pictures of small electronically controlled D'Arsonval meter movement which briefly (0.01 second) interrupts light beam to various galvanometers sequentially thus permitting trace identification.

283. A Bibliography of Information Theory, **F. L. Stumpers.** Massachusetts Institute of Technology, Research Laboratory of Electronics, Cambridge, Mass., Feb. 2, 1953, 46 pp. Over 800 titles arranged by 11 subject classifications.

284. Suggestions for the Decoder Design, **A. K. Susskind.** *D.I.C. 6694, Engineering Memo 14*, Massachusetts Institute of Technology, Servomechanisms Laboratory, Cambridge, Mass., May 22, 1950, 20 pp. Description of three types of decoders converting a 6-binary-digit voltage to an analogue voltage. Detailed comparison of errors and equipment requirements.

285. NACA Oil-Damped V-G Recorder, **I. Taback.** *NACA Technical Note 2194*, National Advisory Committee for Aeronautics, Oct. 1950, 28 pp. New velocity-gravity recorder with improved frequency response, less vibration susceptibility and less field adjustment.

286. Investigations on New Magnetic Recording Media, **M. D. Temple et al.** *Final Report PB-33200*, Brush Development Company, Cleveland, Ohio, Feb. 1944. Final report on project to develop improved magnetic wire by depositing magnetic layers on nonmagnetic material.

287. The Graph-Scope, an Electronic Graph Plotter and Graphical Computer, **A. L. Thomas, Jr.** *Electrical Engineering*, vol. 69, Dec. 1950, pp. 1097–1100. Points are plotted on scope and may be photographed (recorded). Scales can be varied. Equipment description with block diagrams.

288. Design Features of a Magnetic Drum Memory for the National Bureau of Standards Western Automatic Computer (SWAC), **R. Thorensen.** *Proceedings*, Electronic Computer Symposium, May 1952, pp. II: 1–6. Description of co-ordination between magnetic and electrostatic memories to reduce the net access time.

289. High Frequency Effective Incremental Permeability Data for Saturable Magnetic Cores, **W. Titman et al.** Engineering Research Associates, Saint Paul, Minn., May 17, 1951, 24 pp. Approximately 180 "butterfly" curves of 32 cores experimentally determined at 200 kc, 2 megacycles, and 5 megacycles. Four core materials of varying thickness and different convolution were tested by the method described.

290. High-Speed Writing and Selective Altering of Digital Information on Magnetic Surfaces, **D. H. Toth.** *Contract NObsr-42001*, Engineering Research Associates, Saint Paul, Minn., Sept. 1, 1950, 19 pp. Comparison of driving magnetic recording head with hard tubes and thyratrons. Circuits, test procedures, and oscillograms of results are given. With hard tubes, faster rates and higher signal-to-noise ratios were obtained.

291. The Potentialities of Transistors in Digital Computing Circuits, **Dolan H. Toth.** *Contract NObsr-42001*, Engineering Research Associates, Saint Paul, Minn., July 12, 1950, 26 pp. Report gives test results on a number of transistors and considers their use in multivibrators, counters, and blocking oscillators. Circuit diagrams and design methods are given.

292. Preliminary Evaluation of the Harvard-Type Static Magnetic Memory, **Dolan H. Toth.** *Report T.M.-21. XA-19748.* Engineering Research Associates, Saint Paul, Minn., April 12, 1951, 5 pp. Brief description of static magnetic memory and some test results.

293. Binary Counter Uses Two Transistors, **Robert L. Trent.** *Electronics*, vol. 25, July 1952, pp. 100–01.

294. The Effect of Pen-to-Paper Friction in Recording Instruments, **M. J. Tucker.** *Electronic Engineering*, vol. 23, Feb. 1951, pp. 61–63.

295. Information Theory Applied to System Design, **W. G. Tuller.** *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1612–14. Brief paper showing an example of application to air traffic control. Sixteen words are sent to plane.

296. Theoretical Limitations on the Rate of Transmission of Information, **W. G. Tuller.** *Proceedings*, Institute of Radio Engineers, vol. 37, May 1949, pp. 468–78.

297. Digital Recorder, **Arthur W. Tyler.** Eastman Kodak Company, Camera Works, Rochester, N. Y., May 12, 1952, 5 pp. Short proposal for a digital recorder using a storage density of 10^6 bits per square inch. Light from a 2-dimensional array of glow lamps is focused on a rotating photographic disk.

298. Optical and Photographic Storage Techniques, **Arthur W. Tyler.** *Symposium on Large-Scale Digital Calculating Machinery*, Harvard University Press, Cambridge, Mass., 1948, pp. 146–50. Application of photographic storage for digital computers. Estimates of density and access time capability and their limiting factors. Brief consideration of semipermanent storage using phosphors where information is stored by ultraviolet light and is read and erased by infrared light.

299. Review of Input and Output Equipment Used in Computing Systems. *AIEE Special Publication S-53*, "Recording Techniques for Digital Coded Data," **Arthur W. Tyler**, 1953, pp. 3–7. Review of data recording techniques with emphasis on the photographic medium. Values for storage density and rates are given.

300. Commercially Available General-Purpose Electronic Digital Computers of Moderate Price. United States Navy, Office of Naval Research. Washington, D. C., 1952, 37 pp. Collection of seven papers given at a symposium on May 14, 1952. Each paper gives a general description of a computer costing less than $100,000. Included are the JAINCOMP-B1, MONROBOT, CADAC, circle computer, Elecom 100, Consolidated 30-201, and the Miniac.

301. Bibliography of Reports on Wire and Tape Recorders, **United States Office of Technical Services.** *Report SB-7*, United States Department of Commerce, Washington, D. C., March 1949, 8 pp. Eighty-one references to unpublished, unclassified reports available from the government. Primarily captured German documents and war department publications. (This bibliography is cumulative and includes references listed in earlier issues of the bibliography.)

302. Sur une Équation Fonctionnelle de la Théorie de L'Information, **J. Ville.** *Câbles et Transmission*, vol. 5, Jan. 1951, pp. 76–83. Functional equation of information theory. How to estimate quantity of information by given signal source.

303. The Reproduction of Magnetically Recorded Signals, **R. L. Wallace, Jr.** *Bell System Technical Journal*, vol. 30, Oct. 1951, pp. 1145–73. Excellent paper treating quantitatively the losses in the magnetic reproducing process. Experimental and calculated results are compared. Self-demagnetization, spacing loss, thickness loss, gap loss, and losses in the recording and reproducing heads are considered.

304. Analysis and Design of Digital-to-Analog Decoder, **R. L. Walquist.** *M. S. Thesis*, Massachusetts Institute of Technology, Cambridge, Mass., 1951.

305. Magnetic Delay-Line Storage, **An Wang.** *Proceedings*, Institute of Radio Engineers, vol. 39, April 1951, pp. 401–07.

306. Magnetic Triggers, **An Wang.** *Proceedings*, Institute of Radio Engineers, vol. 38, June 1950, pp. 626–69.

307. Miniature Rectifier Computing and Controlling Circuits, **An Wang.** *Proceedings*, Institute of Radio Engineers, vol. 40, Aug. 1952, pp. 931–36.

308. Static Magnetic Storage and Delay Line, **An Wang, W. D. Woo.** *Journal of Applied Physics*, vol. 21, Jan. 1950, pp. 49–54.

309. Development of Cathode-Ray Tube for Film Recording, **S. A. Ward, R. E. Lepic.** *Final Report, Contract W36-039-sc-44573.* National Union Radio Corporation, Orange, N. J., 1951, 66 pp. Details of unsuccessful research to develop a cathode-ray tube which produces an instantaneously visible, permanent record. Desired spot size 0.001 inch and recording time 67 microseconds per element. Electromagnetic radiation and ultraviolet light were investigated.

310. Instrumentation for Flight Testing the XB-45 Airplane, **D. K. Warner.** *Instruments*, vol. 22, April 1949, pp. 314–17. Plane uses photorecorders which photograph 120 dial instruments and recording oscillograph. No data on these devices.

311. The Summation of Different Color Radiations by a Photographic Emulsion, **J. H. Webb.** *Journal*, Optical Society of America, vol. 26, Jan. 1936, pp. 12–23. Verification of van Kreveld's law on the addition of colors applied simultaneously and extension to sequential application. Experimental results with theoretical discussion.

312. Recording System for the Rolling-Ball Flowmeter, **John J. Wedel.** *CIT/JPL Memorandum No. 4-54*, California Institute of Technology, Pasadena, Calif., Dec. 30, 1949, 19 pp. Description of pen recorder used in a specific application.

313. High-Speed Counter Uses Ternary Notation, **Richard Weissman.** *Electronics*, vol. 25, Oct. 1952, pp. 118–21. Describes a simple ternary memory and a fast counter in ternary notation. Ternary counter has high speed and uses fewer tubes than binary counter.

314. Tests on Magnetic Tapes Intended for Operation at 15 Inches per Second (in French), **P. H. Werner.** *Technische Mitteilungen der Schweizer Post-Telegraph-Telephon Verwaltung*, vol. 28, Oct. 1950, pp. 382–88. Also abstracted in *Science Abstracts*, 1951, vol. 54-B, p. 92. Five types of tapes were tested for suitability and degree of interchangeability.

315. A Digital Computer for Scientific Applications, **Charles F. West, J. E. DeTurk.** *Proceedings*, Institute of Radio Engineers, vol. 36, Dec. 1948, pp. 1452–60. Computer description.

316. Studies on Magnetic Recording, **W. K. Westmijze.** *Ph.D. Thesis*, University of Leiden, Leiden, Holland, 1953, 91 pp.

317. Studies on Magnetic Recording, **W. K. Westmijze.** *Philips Research Reports*, vol. 8, June 1953, p. 161.

318. Review of the Present Status of Magnetic Recording Theory, **W. W. Wetzel.** *Audio Engineering*, vol. 31, Nov. 1947, pp. 14–17, 39; vol. 31, Dec. 1947, pp. 12–16, 37; vol. 32, Jan. 1948, pp. 26–30, 46–47. Comprehensive exposition of magnetic recording. Pt. I describes hysteresis

loop tracer and discusses in detail the effects of demagnetization, neutralization, etc., on the hysteresis loop. Pt. II covers erasing, recording, and reproduction in detail using primarily graphical procedures. A-c and d-c bias and erase are compared. Pt. III combines the effects and compares them critically with experimental frequency response curves. Noise and distortion phenomena are examined.

319. A Flux-Sensitive Head for Magnetic Recording Playback, **D. E. Wiegand.** *Proceedings*, Institute of Radio Engineers, vol. 41, March 1953, p. 424. One-paragraph abstract. High-frequency oscillator is used as carrier for signal.

320. An Ultrasonic Memory Unit for the EDSAC, **M. V. Wilkes, W. Renwick.** *Electronic Engineering*, vol. 20, April 1948, pp. 339–45.

321. Position Synchronization of Rotating Drum, **F. C. Williams, J. C. West.** *Proceedings*, Institution of Electrical Engineers, vol. 98, pt. II, Feb. 1951, pp. 29–34.

322. Storage System for Use with Binary Digital Computing Machines, **F. C. Williams, T. Kilburn.** *Proceedings*, Institution of Electrical Engineers, vol. 96, March 1949, pt. III, pp. 81–100; vol. 96, pt. II, April 1949, pp. 183–202.

323. Universal High-Speed Digital Computers: A Small-Scale Experimental Machine, **F. C. Williams, T. Kilburn, G. C. Tootill.** *Proceedings*, Institution of Electrical Engineers, vol. 98, pt. II, Feb. 1951, pp. 13–28. Description and circuits of an experimental serial computer using regenerative cathode-ray tube storage. One appendix describes how information is stored on the tube.

324. Magnetic Recording 1900–1949, **Carmen Wilson.** *John Crerar Library Bibliography Series*, no. 1, John Crerar Library, Chicago, Ill., 1950, 61 pp. Bibliography of 339 abstracts and titles.

325. Progress Report in Miniaturization of Neon Tubes for Digital Recorders. **Gardner P. Wilson, N. R. Gunderson.** *Report TM-562*, United States Naval Ordnance Test Station, Inyokern, Calif., Oct. 9, 1951, 25 pp. Detailed description of experiments and equipment for the development of a small cold-cathode tube for camera use having a striking potential below 90 volts. Mechanical sketches of different tube structures tested. Effect of different gas mixtures and cathode materials. Effect of radio-frequency waves on brilliance.

326. Signal and Noise Levels in Magnetic Tape Recording, **D. E. Wooldridge.** *AIEE Transactions*, vol. 65, 1946, pp. 343–52.

327. A Serial Binary Decoder, **K. N. Wulfsberg.** *Report E-4086*, United States Air Force Cambridge Research Center, Cambridge, Mass., Jan. 1952, 20 pp. Discussion of decoding accuracy and circuit description of 8-digit Shannon decoder.

328. Multichannel Measurement of Physical Effects by Confluent Pulse Technique with Particular Reference to Analysis of Strain, **J. G. Yates et al.** *Paper 1030*, Institution of Electrical Engineers, 13 pp.

329. Data Transmission Equipment, **N. H. Young.** *Proposal 596*, Federal Telecommunication Laboratories, Nutley, N. J., April 6, 1948, 10 pp. Brief proposal mentioning different methods to quantize shaft position.

330. Magnetic Recording of Meter Data, **R. E. Zenner.** *Audio Engineering*, vol. 34, Feb. 1950, pp. 16–17, 33. Usual sound recording equipment used for instrumentation. Discusses errors in magnetic recording quantitatively and methods of avoiding them. Mentioned are amplitude errors, speed variations, and spurious responses.

List of Periodicals

1. *AIEE Transactions*, New York, N. Y.
2. *Acta Physica Austriaca*, Vienna, Austria
3. *Akustische Zeitschrift*, Leipzig, Germany
4. *Annales des Télécommunications*, Paris, France
5. *Archiv der elektrischen Übertragung*, Wiesbaden, Germany
6. *Archiv für technisches Messen*, München, Germany
7. *Arkiv för Matematik, Astronomi och Fysik*, Stockholm, Sweden
8. *Audio Engineering*, Mineola, N. Y.
9. *Aviation Week*, New York, N. Y.
10. *Bell Laboratories Record*, New York, N. Y.
11. *Bell System Technical Journal*, New York, N. Y.
12. *British Journal of Applied Physics*, London, England
13. *Câbles et Transmission*, Paris, France
14. *Chemical Engineering*, New York, N. Y.
15. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, Paris, France
16. *Convention Record*, Institute of Radio Engineers, New York, N. Y.
17. *Electrical Communication*, New York, N. Y.
18. *Electrical Engineering*, New York, N. Y.
19. *Electrical Manufacturing*, New York, N. Y.
20. *Electronic Engineering*, London, England
21. *Electronic Industries and Electronic Instrumentation*, New York, N. Y.
22. *Electronics*, New York, N. Y.
23. *Elektrotechnische Zeitschrift*, Berlin, Germany
24. *Engineer*, London, England
25. *Engineering*, London, England
26. *Funk und Ton*, Berlin, Germany
27. *General Electric Review*, Schenectady, N. Y.
28. *Instruments*, Pittsburgh, Pa.
29. *Interchemical Review*, New York, N. Y.
30. *Journal*, Acoustical Society of America, New York, N. Y.
31. *Journal*, British Institution of Radio Engineers, London, England
32. *Journal*, Franklin Institute, Philadelphia, Pa.
33. *Journal of Applied Physics*, New York, N. Y.
34. *Journal of Research*, United States Bureau of Standards, Washington, D. C.
35. *Journal of Scientific Instruments*, London, England
36. *Journal of the Optical Society of America and Review of Scientific Instruments*, New York, N. Y.
37. *Journal*, Optical Society of America, New York, N. Y.
38. *Journal*, Photographic Society of America, East Pittsburgh, Pa.
39. *Journal*, Society of Motion Picture Engineers, Easton, Pa.
40. *Journal*, Society of Motion Picture and Television Engineers, Easton, Pa.
41. *Journal*, Washington Academy of Sciences, Washington, D. C.
42. *L'Onde Électrique*, Paris, France
43. *Mathematical Tables and Other Aids to Computation*, Washington, D. C.
44. *Mechanical Engineering*, New York, N. Y.
45. *NACA Technical Notes*, National Advisory Committee for Aeronautics, Cleveland, Ohio
46. *Nucleonics*, New York, N. Y.
47. *Optik*, Weimar, Germany
48. *Philips Research Reports*, Eindhoven, Netherlands
49. *Philosophical Magazine*, London, England
50. *Photographic Engineering*
51. *Physica*, The Hague, Netherlands
52. *Physical Review*, New York, N. Y.
53. *Physics Today*, New York, N. Y.
54. *Proceedings*, Electronic Computer Symposium, Institute of Radio Engineers Group on Electronic Computers, Los Angeles, Calif.
55. *Proceedings*, Institution of Electrical Engineers, London, England
56. *Proceedings*, Institute of Radio Engineers, New York, N. Y.
57. *Proceedings*, Institution of Radio Engineers, Sydney, Australia
58. *Proceedings*, National Electronics Conference, Chicago, Ill.
59. *Product Engineering*, New York, N. Y.
60. *RCA Review*, New York, N. Y.
61. *Radio and Television News*, New York, N. Y.
62. *Radio-Electronic Engineering*, New York, N. Y.
63. *Railway Age*, New York, N. Y.
64. *Review of Scientific Instruments*, New York, N. Y.
65. *Siemens Zeitschrift*, Siemensstadt, Germany
66. *Technical Data Digest*, United States Air Force, Dayton, Ohio
67. *Technische Mitteilungen der Schweizer Post-Telegram-Telephon Verwaltung*, Switzerland
68. *Tele-Tech*, New York, N. Y.
69. *VDE-Fachberichte*, Berlin, Germany
70. *Wireless World*, London, England
71. *Zeitschrift für wissenschaftliche Photographie, Photophysik und Photochemie*, Leipzig, Germany

No Discussion

A New Portable Telegraph Transmission Measuring Set

S. I. CORY
MEMBER AIEE

THE provision of good telegraph service requires that the circuits be designed, engineered, and maintained so as to afford good transmission quality. The measurement of a single quantity, known as telegraph distortion,[1] at the receiving end is sufficient to establish the transmission quality since it contains all signal degrading effects, such as interference, poor relay performance, and wave-shape distortion. Therefore, the ability to measure distortion conveniently and accurately is of prime importance.

Suitable distortion-measuring devices are provided in the Bell System for use at telegraph test boards in the larger offices, in maintenance work, and in special investigations in the field,[3,4] as well as for development work. However, no device has been available which fulfills satisfactorily the requirements for use at customers' stations and in small central offices.

This gap is being filled by a set of new design especially suited to this use. This paper describes the field-trial model of this new design.

The Requirements: How They Are Met

Because this set is to be carried by the station maintenance men to customer stations along with other gear, an important objective in its design was to build as small and lightweight a set as is consistent with good performance. Another requirement was that the set must measure on a start-stop basis since Bell System circuits are now generally operated by start-stop teletypewriters.[2] These requirements are met by an all-electronic design which provides an improved type of indication on a small cathode-ray tube and which uses standard miniature parts and tubes, and, for the most part, conventional present-day circuitry. The design of this model is centered around a 1-inch cathode-ray tube and an associated optical magnifier for the distortion indicator. This, along with the use of multivibrator timing circuits, requires only a small power supply which is also contained in the set. Fig. 1 shows a view of this field-trial model which is 6 inches high by $3\frac{1}{2}$ inches wide by $6\frac{5}{8}$ inches deep and weighs only about 6 pounds, complete with carrying case. For practical reasons of manufacturing economy, ease of maintenance, etc., the final design for production will contain a 2-inch cathode-ray tube without the optical magnifier, and it will be somewhat larger and heavier than the trial model.

In line with the general Bell System requirements for a suitable measuring set, it is designed to measure teletypewriter and teletypesetter signals in working circuits of any type without disturbing the service. Also it may be used in any circuit encountered in portable use, or at testboards in central offices. Its input is a 100-ohm nonreactive resistance which is inserted in series in the local circuit and has practically no effect on the circuit under test. The set may be used to indicate the distortion of miscellaneous or recurring 5- or 6-unit code teletypewriter signals at 60, 75, or 100 words per minute by making simple adjustments. A simple and quick local check procedure is followed which does not necessarily require a special source of undistorted signals, i.e., it may be made using the signals received from the distant station, a feature which is of considerable advantage in portable use. The set requires only plug connections to a source of 110-volt 60-cycle power and to the circuit to be measured.

A new method of displaying distortion on a cathode-ray tube is employed to improve the ease of use as compared to other designs employing cathode-ray tubes as indicators.[4-6] Two types of indication are used—the pip and peak indications shown in Fig. 2. In the pip indication of distortion the instantaneous value of the displacement of each signal transition is indicated by means of easily observed "pips" on a 0-to-50-per-cent distortion scale along a horizontal axis on the face of the tube. The pip is upwards from the horizontal trace for space-to-mark transitions and downward from the trace for mark-to-space transitions. The pips have bright trailing edges or tails which extend to the right or left depending on the sign of the displacement. Accordingly, they are easy to observe. The peak indication of distortion is a persisting indication of the maximum displacement, otherwise known as the peak value of the total distortion, using the position of a spot of light on the distortion scale. The pips are also superposed on the spot of light and appear in a vertical line through the center of the spot, to advise the observer whether or not signals are being received. This type of indication is comparable to the meter indication of a set now in general use in large central offices.[3] It is easy to observe, even if the light is poor, as may often be the case in portable use. Accuracy of indication compares well with that of larger and more expensive sets, it being in the order of ±2-per-cent distortion.

Status

The model sets have been given extensive trials in the Bell System with satisfactory results. Also, it is understood that the United States Army Signal Corps has had satisfactory experience with the models and they believe that this type of set will be valuable to the Armed Forces. Final development and design of the circuit and equipment is well along, leading to the production of sets which will be coded "164C1 Telegraph Transmission Measuring Set." The commercial product will be somewhat larger and heavier than the trial model.

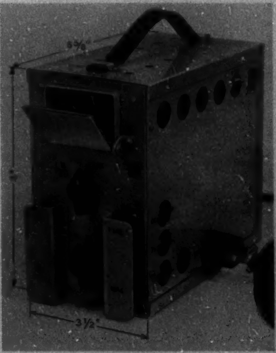


Fig. 1. Oblique front view of the field-trial model set

Paper 54-105, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 21, 1953; made available for printing December 1, 1953.

S. I. CORY is with the Bell Telephone Laboratories, New York, N. Y.

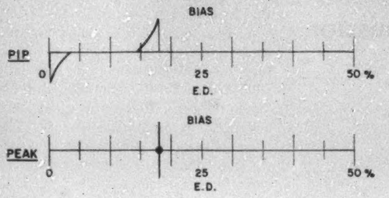


Fig. 2. Pip and peak indications

Description

Any start-stop telegraph transmission measuring set contains certain fundamental parts: means for receiving the signals, means for timing the duration of characters and pulses, and means for indicating the displacement of the transitions from their proper positions, i.e., the distortion. In this model set the means for receiving signals is the trigger-tube input circuit shown at the left in Fig. 4. This operates the character timer and the pulse timer which are both multivibrator circuits. The output of the pulse timer contains an integrating circuit and amplifier which produces a triangular-shaped output for the horizontal sweep of the cathode-ray tube when pips are being observed. This provides the time base for the distortion scale. The beam is at the left, or zero-distortion end of the scale, at the instants that transitions occur in perfect signals and at the right, or 50-per-cent distortion end of the scale, halfway between these instants. A differentiating circuit operating from the input circuit produces a pip at the time of occurrence of each transition in the signal and these pips are applied to the vertical plates of the cathode-ray tube to provide the indication of instantaneous values of distortions. Peak-distortion indication requires that the triangular-shaped voltage wave be sampled at the time of occurrence of each transition in the signal and the sample voltage applied to a storage capacitor in the peak voltmeter circuit. The stored voltage controls the position of the spot of light along the distortion scale to provide a persisting indication of the peak value of the distortion.

The operation of the circuit will now be described in more detail, using Figs. 3 and 4 together.

Input Circuit

The telegraph signals are impressed on a 100-ohm input resistance *R* which is inserted directly into the circuit to be measured. The input tube *V1* is a double triode used in a bistable trigger circuit whose cathode circuit contains the input resistance *R*. Practically square-wave signals are obtained at the output of *V1* regardless of the waveshape of the input signals. Its performance over the operating range is similar to that of a receiving relay. The output of tube *V1* for undistorted (perfectly timed) input signals is indicated at *B*.

Character Timer

The output of tube *V1* is impressed upon the character timer which is a 1-shot multivibrator using tube *V2*. Character timing begins upon the occurrence of the start pulse and ends at the beginning of the stop pulse as is indicated at *C*. Character timer adjustment is made by means of the CHAR potentiometer.

Pulse Timer

The starting and stopping of the pulse timer is controlled by the character timer. At the beginning of the character-timing interval the pulse timer tube *V3* is started and operates as a free-running multivibrator during the character, producing a practically square-wave output with pulses at twice the signalling frequency. Thus, the action of tube *V3* is similar to that of a start-stop oscillator. Its output is indicated at *D*. Speed adjustment is made by means of the SPEED potentiometer. There are three of these, one for each speed.

Integrating Circuit and Amplifier

The output of the pulse timer is impressed upon an integrating circuit having a large time constant so that the volt-

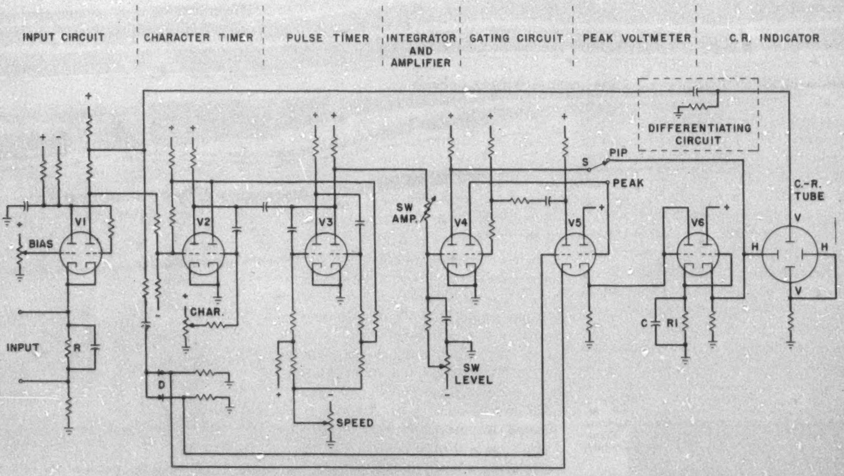


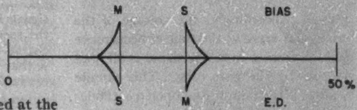
Fig. 3. Functional schematic. Ground shown is to chassis

age change on the capacitor of the integrating circuit is only a few volts during each half-cycle of the pulse timer output. Because of this small change, the variation in voltage is practically linear with time. This linear change is amplified by the left portion of tube *V4* to obtain the linear sweep voltage indicated at *E*. This sweep voltage is applied to the horizontal plates of the cathode-ray tube when the switch *S* is operated to PIP. It causes the beam to travel from 0 to 50 per cent on the distortion scale during the first half-cycle of the pulse timer and from 50 per cent to 0 during the second half-cycle. SW AMP and SW LEVEL potentiometers are provided for the proper adjustment of the horizontal sweep as viewed on the *CR* tube.

Differentiating Circuit and Pip Indication

The differentiated output of tube *V1* is applied to the vertical plates of the cathode-ray tube to produce a pip for each transition in the signal. As the beam moves back and forth along the horizontal axis of the tube in a linear fashion for each unit part of the signal, it is momentarily deflected upward or downward depending on whether the transition is space-to-mark or mark-to-space. Pips are produced along the horizontal trace as shown at *F*, indicating the time of occurrence of the transitions. If the transitions occur at their proper positions (zero distortion), as shown at *F*, they are all located at the extreme left, or zero-distortion end of the scale. Distortions of various amounts will be indicated by the pips occurring at other positions along the 0-to-50-per-cent scale depending upon the amount of distortion. The time-constant of the decay of the pip voltage is adjusted so that a bright and easily observed tail is produced in the direction in which the spot is moving in response to the horizontal sweep. Thus, if a transition occurs while the spot is moving towards zero per cent on the scale the transition occurs ahead of the correct time. In this case a pip having a tail to the left will be produced, indicating marking bias[7] if the pip is upwards and spacing end distortion[7] if the pip is downwards. If a transition occurs while the spot is moving towards 50 per cent on the scale, the transition occurs later than it should. In this case a pip having a tail to the right will result, indicating spacing bias if the pip is upwards and marking end distortion if the pip is downwards. A legend showing this, which is printed on the front of the set for the guidance of the user, is shown in Fig. 5. The distortion is read as that indicated at the time of the sudden upwards or downwards change in spot position. Bias is the average position of the upward pips since bias is, by definition, the average displacement of space-to-mark transitions.

Fig. 5. Legend indicating type and sign of distortion



Peak Distortion Circuit and Indication

The persisting indication of the peak value of distortion is obtained by causing each transition in the signal to open a gate valve, the right portion of tube *V4*, which in its closed condition suppresses the triangular voltage wave produced by the left portion of *V4*. The gating pulses are obtained by rectifying the differentiated output of tube *V1* by means of the diodes at *D*. The right portion of tube *V4* is rendered nonconducting by the gating

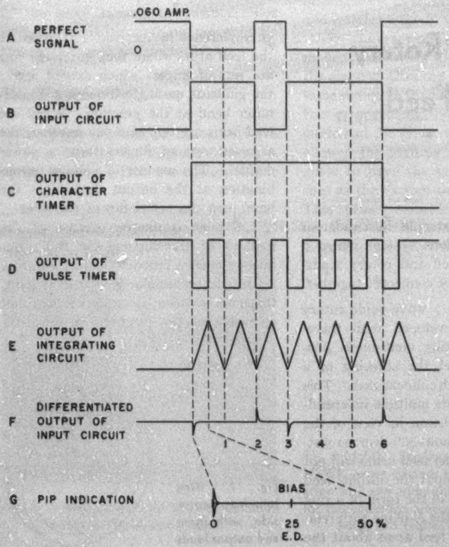


Fig. 4 (left). Wave shapes

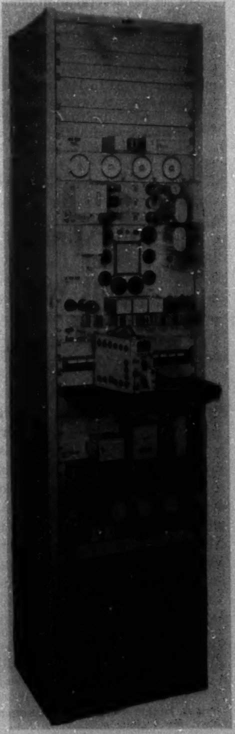


Fig. 6 (right). Model set on the shelf of the 118C3 telegraph transmission measuring set

pulses. Accordingly, the voltage of the triangular wave at the time of occurrence of each transition in the signal is allowed to pass by the gate tube. This sample voltage is then impressed upon a cathode-follower tube, the right portion of tube *V5*, and then upon a peak-voltmeter circuit containing a storage capacitor *C* and the left portion of *V6*, which is used as a diode. The output of the peak-voltmeter circuit is impressed upon the horizontal plates of the cathode-ray tube through a cathode follower, the right portion of *V6*. This causes the beam to be deflected by an amount corresponding to the peak value of distortion. The storage capacitor is slowly discharged by high resistance *R1* and the beam returns slowly towards zero if peak distortions of the same amount or larger do not follow immediately. It may be returned to zero quickly by operating a reset button which connects a low resistance across the storage capacitor.

OTHER FEATURES

The set has several additional interesting features which should be pointed out:

1. It may be used at any speed between approximately 55 to 110 words per minute since this range is covered by each of the three speed-controlling potentiometers. Also, it may be used with 5- or 6-unit code signals having any speed within the range of the speed controls by merely adjusting the character and pulse-timing potentiometers so that the timers operate properly with the received signals.

2. Three posts are provided at the rear of the set to which an external oscilloscope of any desired size may be connected to serve as an external distortion indicator. This is of some value in obtaining an enlarged indication of distortion for use in demonstrations or even in the normal use of the set in the case of certain installations.

3. An adjustable mirror is provided on the front of the set for the convenience of the user; see Fig. 1. By properly adjusting the position of the mirror the set may be used horizontally on a table with the observer looking down into the mirror, or it may be mounted vertically, as at a telegraph board position, in which case the mirror may be adjusted by each observer to accommodate his eye-level. The set may also be used horizontally at eye-level or tilted properly if below eye-level in which case the mirror is not used, i.e., the observer looks directly into the hood.

4. Using the three posts on the rear of the set and an external cathode-follower of sufficient power capacity to operate a recording meter, a record may be made of the variation in the peak distortion with time. Such a circuit should, of course, have an automatic reset feature which will restore the peak indication to zero periodically.

Conclusions

This new set is a striking example of what can be done in reducing size, weight, and cost using a new approach in circuit design and a moderate amount of miniaturization. The trial model set is but a small fraction of the size and weight of the 118-type set[3] which is now generally used in the larger offices and yet it provides comparable features. An idea of the relative sizes of the two sets is afforded in Fig. 6 in which the model set is resting on the writing shelf of the 118-type set.

References

1. MEASUREMENT OF TELEGRAPH TRANSMISSION, **H. Nyquist, R. B. Shanck, S. I. Cory.** *AIEE Transactions*, vol. 46, Feb. 1927, pp. 367–76.

2. FUNDAMENTALS OF TELETYPEWRITERS USED IN THE BELL SYSTEM, **E. F. Watson.** *Bell System Technical Journal*, New York, N. Y., Oct. 1938.

3. RECENT DEVELOPMENTS IN THE MEASUREMENT OF TELEGRAPH TRANSMISSION, **R. B. Shanck, F. A. Cowan, S. I. Cory.** *Bell System Technical Journal*, New York, N. Y., Jan. 1939, p. 143.

4. A CATHODE-RAY TELEGRAPH DISTORTION MEASURING SET, **W. T. Rea.** *Bell Laboratories Record*, New York, N. Y., April 1947.

5. AN ELECTRONIC DISTORTION MEASURING SET FOR START-STOP TELEGRAPH SIGNALS, **L. K. Wheller, R. S. Tissington,** *The Post Office Electrical Engineers' Journal*, London, England, April 1950.

6. AN IMPROVED T. D. M. S. G. T., **Dain.** *A. T. E. Journal*, Jan. 1953.

7. EFFECT OF TELEGRAPH DISTORTION ON THE MARGINS OF OPERATION OF START-STOP RECEIVERS, **W. T. Rea.** *Bell System Technical Journal*, New York, N. Y., July 1944.

◆

No Discussion

An Annular Wave-Guide Rotary Joint with Wave-Guide Feed

LOUIS D. BREETZ
NONMEMBER AIEE

THERE are many engineering problems in the field of electronics which require a good method of transmitting radio frequency energy from a fixed unit to a rotating unit. For radio frequencies at which wave guides are normally employed, the rotary joint herewith described is especially suitable, particularly where there is a multiplicity of independent transmission lines involved in a relative motion; for example, in a radar or radio communications system where a combination of fixed and rotary transmission lines must be connected together.

For this purpose a wave-guide rotary joint has been developed with wave-guide feed permitting continuous 360-degree rotation and the stacking of a number of joints with colinear axes. This type of joint permits multiple independent channel transmission in view of the fact that the continuous 360-degree rotating output (or input) feed arms will not interfere with the input (or output) feed arms. The joint is in the form of a ring wave guide bent in the *E*-plane with relative motion of the feed arms about the circumference of the ring. Attached to the feed arms where they enter the ring are sets of fingers which extend across the guide at an angle to form a double-miter bend at the junction. When the feed arms are opposite one another, the adjacent sets of fingers form a power divider. The manner of voltage recombination at the output determines the band-pass characteristics of the joint.

The joint is unique because of the method of transferring the radio frequency energy into and out of the ring guide. In an annular guide rotary joint, the main problem is to avoid improper operation at the positions of crossover

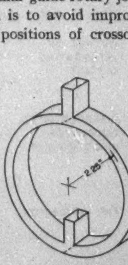


Fig. 1. Ring joint from narrow side, with input and output feeds

---

Paper 54-179, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted April 30, 1953; made available for printing November 30, 1953.

LOUIS D. BREETZ is with the Naval Research Laboratory, Washington, D. C.

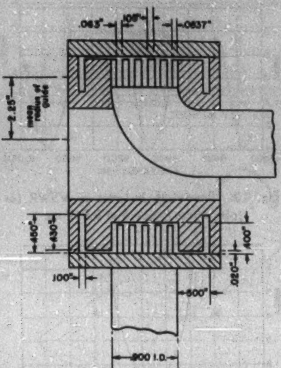


Fig. 2. Ring joint from wide side, showing the fingers and chokes. The inner ring with its fingers, chokes, and feed line rotates with respect to the outer ring and its feed line and fingers

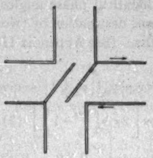


Fig. 3 (left). The fingers from the narrow side of guide in position of crossover arrows show relative movement

or where the input and output lines are in the same immediate vicinity. For limited bandwidths and low power levels this problem was satisfactorily met.

The Wave Guide Rotary Joint

The joint is made so that the wave guide is bent in the *E*-plane forming a complete ring. The wave-guide input and output junctions are at the wide sides of the ring guide. At the junctions are sets of fingers directed at such an angle as to make *E*-plane bends for the feed lines going into the ring guide. The fingers on input and output guides are made so they can interleave, thus permitting 360-degree rotation. The bearing surfaces of rotation, which are at the edges of the wide side of the ring guide, are at regions of high current flow and hence chokes are required at the corners of the outer wide side.

The fingers of the *E*-plane double-miter bends are set at about a 45-degree angle with the tangent to the circumference where the fingers enter the ring. The mean path length of the miter bend is nearly $\lambda_g/4$. The exact angle of the fingers must be determined experimentally for the value that gives the least reflection at crossover, since such things as capacity between the sets of fingers and the possibility of the fingers not being parallel when fully meshed are suspected of having an effect on the operation.

The model joint (Figs. 1 through 5) was made to perform at approximately 9,000 megacycles. Using a guide of 0.90 inch by 0.40 inch, inside dimensions, it was found that the fingers 0.063 inch wide, consisting of six on one set and five on the other with 0.105-inch separation between adjacent fingers, served satisfactorily. A distance of 0.03 inch between the opposite side walls and the ends of the sets of fingers was sufficient. This arrangement gave satisfactory screening and interleaving.

The chokes were $\lambda/2$ deep and were made to bend with the ring guide around the circle so that the chokes were on the same radius as the bent guide. Through this arrangement the chokes and wave guide had the same circumference, which allowed the exciting waves of the wave guide to have the same phase relationship as the excited waves in the chokes. This phase coincidence cannot easily occur in an *H*-plane-bend ring (*H* is magnetic intensity) where the inner and outer chokes and the wave guide would have different circumferences. A model of the latter type was made and it was shown that the chokes and ring guide of unlike circumference actually gave inferior performance of the joint as against an *H*-plane-bend ring without any chokes. The difference was found in the measurement of joint input impedance with rotation. (See Fig. 6.)

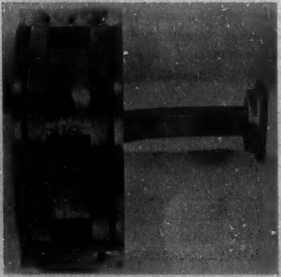


Fig. 4. Inside view of joint showing construction of fingers and chokes

Fig. 5. Outside view of rotary joint showing feed lines

The mean circumference of the ring guide is such that the mean path length F_1 to F_4 and mean path length F_5 to F_8 (Fig. 7) have a difference equal to an integral number of wave-guide wave lengths for the center operating frequency. Points F_1 through F_8 are mid-points in the cross sections of transmitting channels within the joint, and lines connecting these points are mean path lengths of transmission. In the crossover region, or when entrant and exit guides are overlapping (the mid-position is sketched in Fig. 7) the fingers act as power and voltage dividers with power and voltage divided as $d_1/(d_1+d_5)$ and $d_5/(d_1+d_5)$, where d_1 through d_5 are lengths in the *E*-plane dimension of the cross section.

The difference in the path lengths from F_1 to F_4 and from F_5 to F_8 varies only slightly with degree of rotation at crossover. For example, in the described model the difference in lengths from F_1

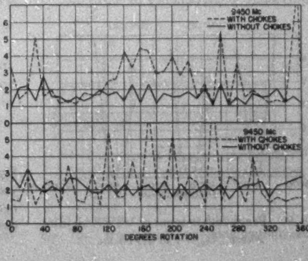


Fig. 6 (right). H-plane-bend ring joint: VSWR versus degrees rotation for operation with and without half-wave chokes for the condition where the radii of the guide and chokes are unequal

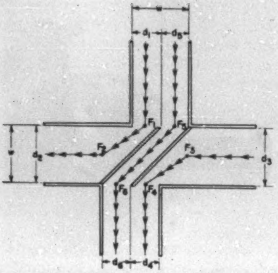


Fig. 7. Cross-section view of narrow side of ring guide showing mean path lengths at crossover

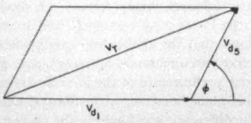


Fig. 8. Funicular diagram of voltages at output of joint

to F_4 and from F_5 to F_6 varies with d_1 according to the following tabulation.

d_1, inches	F_1 to F_4–F_5 to F_6, inches
0.4	13.71
0.3	13.68
0.2	13.66
0.1	13.65

The path length difference varies only 0.4 per cent. (See Appendix I for derivation.) In the model joint, the radius of the annular guide is 2.25 inches; the mean circumference is 14.1 inches, and the path-length F_5 to F_6 is 0.56 inch. These dimensions give a path difference of 1.66 inches, or 34.69 centimeters (cm) at mid-position. This path difference is exactly $7\lambda_g$ for a wave-guide wave length of 4.954 cm ($\lambda=3.36$ cm in air) which is therefore a center-operating wave length equivalent to a frequency of 8,918 mc.

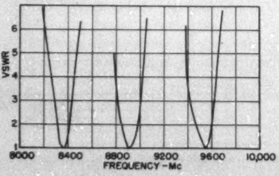


Fig. 9. Bandwidth in terms of VSWR for joint where $N\lambda_{g_0}=7$ at 8,918 mc

Bandpass Problem

At any noncrossover position, bandpass problems do not exist except those introduced by the chokes and wave guide. At crossover position the arrangement forms a power and voltage divider in which P_1 and P_5, V_1 and V_5 are powers and voltage magnitudes corresponding to the two wave-guide paths whose *E*-plane dimensions are denoted by d_1 and d_5 in Fig. 7. The power and voltage divide as follows: $P_1/P_5=V_1/V_5=d_1/d_5$. Assuming negligible dissipation losses, the voltage magnitudes in the two paths at the place of recombination are the same as at the place of division, for the two paths. Also the real power entering the joint is equal to the real power leaving the joint. Whenever the voltages from the two paths at the place of recombination are of unlike phase which is always the case except when the difference in length of the two paths is an integral number of guide wave lengths, the maximum possible transfer of real power through the joint is not obtainable. Any frequency change results in a phase change, and therefore a change in real power output. The power that is not accepted by the joint can be likened to power that is reflected at the entrance to the joint. If one assumes that there are no other reflections, the bandpass characteristic is dependent upon the way the phases of the voltages over the two paths recombine at the output as the guide wave length varies.

The resultant voltage at the output is given by

$$|V_T|^2=|V_{d_1}|^2+|V_{d_5}|^2+2|V_{d_1}||V_{d_5}|\cos\phi \quad (1)$$

where

$|V_T|$ = magnitude of the output voltage
$|V_{d_1}|$ = magnitude of voltage from path whose *E*-dimension is d_1 (Fig. 7)
$|V_{d_5}|$ = magnitude of voltage from path whose *E*-dimension in Fig. 7 is d_5
ϕ = phase difference angle of V_{d_1} and V_{d_5}

Fig. 8 shows a funicular diagram of these voltages.

This equation may also be presented in terms of the phase-shift difference angle

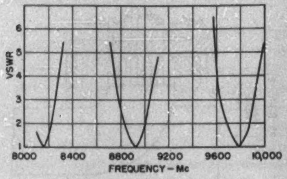


Fig. 10. Bandwidth in terms of VSWR for joint where $N\lambda_{g_0}=5$ at 8,918 mc

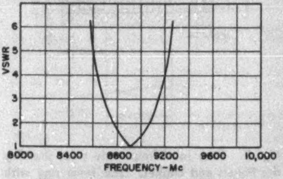


Fig. 11. Bandwidth in terms of VSWR for joint where $N\lambda_{g_0}=3$ at 8,918 mc

of two waves originating from the same plane with nearly identical phase angles but reaching the same destination by two different path lengths. (See Appendix II for derivation.)

$$|V_T|^2=|V_{d_1}|^2+|V_{d_5}|^2+2|V_{d_1}||V_{d_5}| \cos\left[2\pi N\left(\frac{\lambda_{g_0}}{\lambda_g}-1\right)\right] \quad (2)$$

The power accepted $=P_a/P_i=|V_T|^2$ where incident power (P_i) to the joint is considered as unit power and incident voltage (V_i) is considered as unit voltage corresponding to unit incident power, and where $|V_{d_1}|+|V_{d_5}|=|V_i|$. Some acceptable value for P_a/P_i determines bandwidth, i.e., semibandwidth is

$$\Delta\lambda_g=\lambda_{g_0}\Big]_{\frac{Pa}{Pi}=1}-\lambda_g\Big]_{\frac{Pa}{Pi}} \quad (3)$$

where $P_a/P_i=$ the right-hand member of equation 2.

Since no power output is obtainable when two oppositely phased voltages are of equal magnitude, the worst case for consideration of bandwidth at crossover is when $d_1=d_5$, as shown in Fig. 7, and

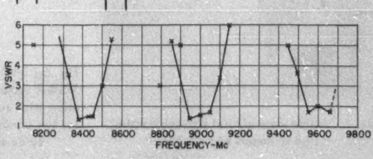


Fig. 12. Measured VSWR versus frequency at worst crossover position for joint whose path-length difference is 7λ for 8,918 mc

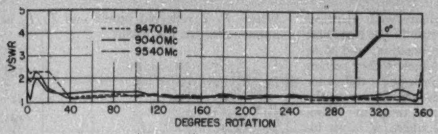


Fig. 13 (above). VSWR versus degrees rotation at three frequencies

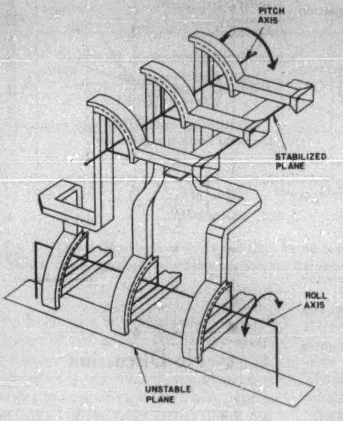


Fig. 14 (left). Stabilization axis joint

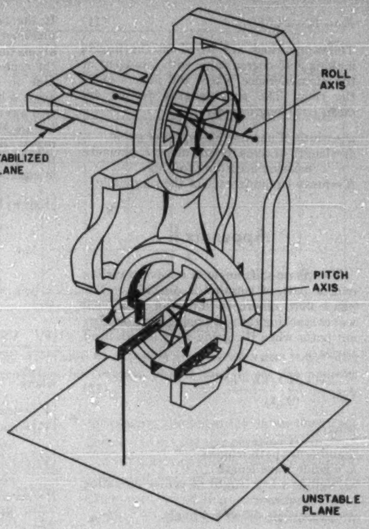


Fig. 15 (right). Stabilization axis joint

where $|V_{d_1}|=|V_{d_2}|=0.5$.

For the worst position

$$\frac{P_a}{P_t}=0.5+0.5\cos\left[2\pi N\left(\frac{\lambda_{g_0}}{\lambda_g}-1\right)\right] \qquad (4)$$

The guide wave length corresponding to the acceptable value of $P_a/P_t=K$ from equation (4) is

$$\lambda_g=\frac{\lambda_{g_0}}{\dfrac{\cos^{-1}(2K-1)}{2\pi N}+1} \qquad (5)$$

Bandwidth is also expressible in terms of voltage standing wave ratio (VSWR) set up in a transmission line by the joint. If R is the permissible VSWR at the band-limiting frequency, then (See Appendix III for derivation.)

$$\lambda_g=\frac{\lambda_{g_0}}{\dfrac{\cos^{-1}\left\{2\left[\dfrac{4R}{(R+1)^2}\right]-1\right\}}{2\pi N}+1}=\frac{\lambda_{g_0}}{\dfrac{\cos^{-1}\left\{\dfrac{8R}{(R+1)^2}-1\right\}}{2\pi N}+1} \qquad (6)$$

Measurements

Figs. 9, 10, and 11 show how bandwidth at worst crossover position of the feed line, in terms of VSWR each side of a center operating frequency, varies for different values of $N\lambda_{g_0}$, as calculated by the equation 6.

Fig. 12 shows a plot of actual measured maximum VSWR versus frequency for the model whose ring circumference was 14.1 inches and whose path length difference (Fig. 7) F_1 to F_4 minus F_5 to F_6, was $7\lambda_g$ for 8,918 megacycles (mc). This generally follows the calculated values plotted in Fig. 9. Fig. 13 shows a plot of VSWR versus rotation for three spot frequencies. The operation is seen to be satisfactory for receiving purposes for the entire 360-degree rotation. At noncrossover the joint is not frequency sensitive, and frequency limitations become those of the chokes and the guide itself.

Although bandwidths are not extremely great when 360-degree rotation is required, the joint may prove useful when the bandwidth need be no wider than several hundred megacycles in the vicinity of 9,000 mc. Receiving losses were about 0.2 decibels for noncrossover and 0.5 decibels for crossover within the passband.

At transmitter power levels the joint model handled about 200 kilowatts at noncrossover positions and 50 kilowatts at crossover before any evidence of voltage breakdown occurred.

Other Applications

Figs. 14 and 15 show arrangements where the principle involved in the joint can be adapted to use in a stabilized-to-unstabilized-plane coupling for multiple-transmission lines. It should be broad band since no crossover position is anticipated. Fig. 16 suggests an arrangement for a wave-guide switch.

Appendix I

Referring to Fig. 7, and assuming that the electrical path lengths closely follow the mean geometrical paths, the path lengths at crossover are as follows:

$$F_1 \text{ to } F_2=\sqrt{\left(\frac{w}{2}\right)^2+\left(w-\frac{d_1}{2}\right)^2} \qquad (7)$$

$$F_2 \text{ to } F_3=K-2w+d_1 \qquad (8)$$

$$F_3 \text{ to } F_4=\sqrt{\left(\frac{w}{2}\right)^2+\left(w-\frac{d_1}{2}\right)^2} \qquad (9)$$

$$F_1 \text{ to } F_4=2\sqrt{\left(\frac{w}{2}\right)^2+\left(w-\frac{d_1}{2}\right)^2}+K-2w+d_1 \qquad (10)$$

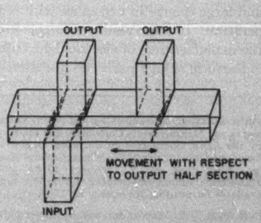


Fig. 16. Wave-guide switching arrangement

$$F_3 \text{ to } F_8 = w\sqrt{2} \qquad (11)$$

These equations are based on the fingers making a 45-degree angle with a tangent at the circumference where the fingers enter the ring. The path lengths defined are mean path lengths where

w = width of narrow dimension of guide
d_1 = length in cross section along E-plane dimension according to Fig. 7
K = mean circumference of the ring guide

Appendix II

The phase-shift angle of a wave traveling over a path of length l is $\phi = 2\pi l/\lambda$. The phase-shift difference angle of two such waves reaching the same point by two different paths when the guide wave length itself varies is

$$\phi = \frac{2\pi(\lambda_{g_0} - \lambda_g)(l_1 - l_2)}{\lambda_{g_0}\lambda_g} \qquad (12)$$

where

l_1 and l_2 are path lengths
λ_g = guide wave length
λ_{g_0} = guide wave length at center operating frequency
ϕ = phase-shift difference angle.

If the two waves originate with identical phase angles, as is the case in the ring joint at maximum crossover position, and if at the center operating frequency they recombine in phase, the phase-shift difference angle ϕ is also the phase angle between the two waves at the output of the joint. Also, $l_1 - l_2 = N\lambda_{g_0}$ = constant, where N is an integer. Substituting $N\lambda_{g_0}$ for $l_1 - l_2$ in equation 12 and then substituting this equation in equation 1

$$|V_T|^2 = |V_{d_1}|^2 + |V_{d_2}|^2 + 2|V_{d_1}||V_{d_2}| \cos\left[2\pi N\left(\frac{\lambda_{g_0}}{\lambda_g} - 1\right)\right] \qquad (13)$$

Appendix III

$$\left|\frac{V_r}{V_i}\right| = \frac{\text{VSWR} - 1}{\text{VSWR} + 1} = \frac{R-1}{R+1} \qquad (14)$$

where

$|V_r|$ = magnitude of reflected voltage
$|V_i|$ = magnitude of incident voltage = unit voltage
$|V_r|/|V_i|$ = magnitude of voltage-reflection coefficient
R = permissible VSWR at the band-limiting frequency

Where P_r is reflected power and P_r/P_i is power reflection coefficient

$$\frac{P_r}{P_i} = 1 - \frac{P_a}{P_i} \qquad (15)$$

$$\frac{P_a}{P_i} = 1 - \frac{P_r}{P_i} = K \qquad (16)$$

$$\frac{P_r}{P_i} = \frac{|V_r|^2}{|V_i|^2} \qquad (17)$$

and from equations 14, 16, and 17

$$K = 1 - \left(\frac{R-1}{R+1}\right)^2 = \frac{4R}{(R+1)^2} \qquad (18)$$

Substituting equation 18 into equation 5, the band limits are given by

$$\lambda_g = \frac{\lambda_{g_0}}{\dfrac{\cos^{-1}\left\{2\left(\dfrac{4R}{(R+1)^2}\right) - 1\right\}}{2\pi N} + 1} = \frac{\lambda_{g_0}}{\dfrac{\cos^{-1}\left\{\dfrac{8R}{(R+1)^2} - 1\right\}}{2\pi N} + 1} \qquad (19)$$

◆

No Discussion

A Tungsten Resistance Thermometer

F. R. SIAS
ASSOCIATE MEMBER AIEE

J. R. MACINTYRE
MEMBER AIEE

A. HANSEN, JR.
MEMBER AIEE

Suitability of Tungsten

THE use of tungsten wire for the sensing element of a resistance thermometer is not new to the art. However, to the best knowledge of the authors, this thermometer, Fig. 1, is the first successful one with stabilities as good as, or better than, platinum.

Tungsten offers many advantages as the sensing element of a resistance thermometer. Among these are its high melting point, 3,380 degrees centigrade (C), and large change in resistance with temperature. It is also readily available from the manufacturers of electric light bulbs, in the form of high-purity ductile wire.

In processing the tungsten wire for use in the resistance thermometer, use is made of the fact that its temperature coefficient of resistance is a function[1] of the annealing temperature. High-quality resistance thermometers using other metals generally depend upon a costly selection process to obtain wire with the precise coefficient necessary to maintain the desired calibration accuracy. The temperature coefficient which has been selected for the tungsten thermometer is obtained by an annealing temperature well outside its operating range, in order to prevent further changes during normal use.

The linearity of the temperature-resistance curve is also an important factor in a resistance thermometer. Fig. 2 shows the departure from linearity of tungsten as compared to several other metals commonly used. The curves were determined by assuming a linear relationship between the 0 and the 100 C points and then noting the departure from this arbitrary relationship at all other points. The departure from linearity has been converted from resistance to degrees centigrade for convenience. It will be noted that the linearity of tungsten is within the range of materials now being used.

Tungsten has to be protected by an inert atmosphere in order to prevent reaction with the surrounding media. This fact does not prove to be a disadvantage, since it is necessary to hermetically seal the bulb for low-temperature operation, and the use of helium as the inert gas improves the thermal response of the bulb. Helium gas also has the property of not diffusing through the metal wall.[2] The loss of helium through the glass of the seal is negligible because of the small area involved.

Description

Fig. 3 illustrates the mechanical construction of the thermometer. The coiled

Fig 1. The tungsten resistance thermometer

Paper 54-16, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted October 13, 1953; made available for printing November 12, 1953.

F. R. SIAS, J. R. MACINTYRE, and A. HANSEN, JR., are with the General Electric Company, Lynn, Mass.

Vo
Pt.
19

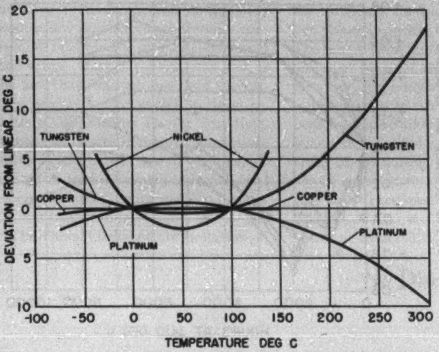


Fig. 2. Curves showing linearity of resistance versus temperature as referred to a theoretical straight-line relationship between 0 and 100 C

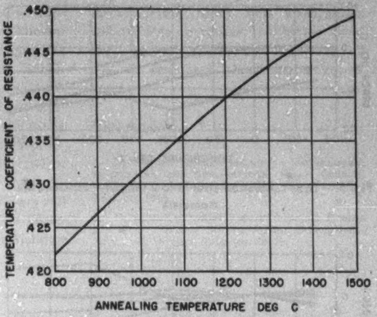


Fig. 4. Curve of annealing temperature versus temperature coefficient for tungsten wire as used in the thermometer

tungsten wire is wound onto a ceramic form with nickel leads attached by welded "wrap-around" pieces as shown. The tungsten is attached to the nickel lead wires by a special welding technique to produce low-resistance high-stability, connections. The third lead of the 3-wire bulb is welded to one of the other leads as close to the ceramic form as feasible.

The nickel leads are then cut to the desired length and welded to the hermetic-seal leads. This whole assembly is then placed into the head and tube assembly, and the hermetic seal soldered or brazed to the brass head. The tungsten-wound ceramic form is spaced from the inside wall of the tubing by means of projections at each end of the ceramic piece. The thermometer is now ready to be calibrated and adjusted.

In the manufacture of a resistance thermometer three qualities of the resistance element must be established. These are:

1. The resistance-temperature coefficient;
2. The resistance at some reference temperature;
3. The stability.

The first adjustment to the bulb is made by placing the assembly in an annealing furnace to set the temperature coefficient.

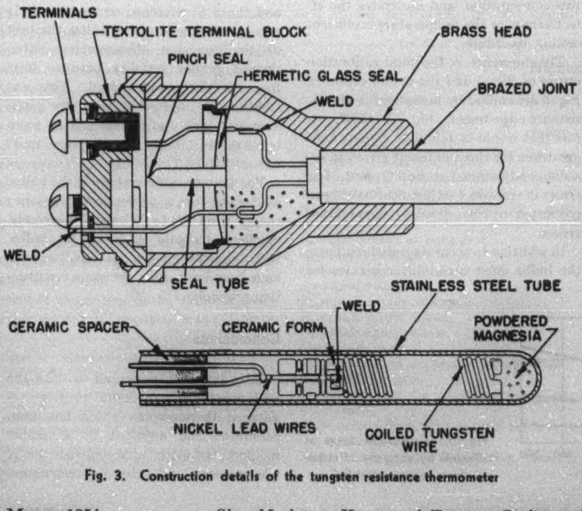


Fig. 3. Construction details of the tungsten resistance thermometer

Smithells[1] has shown that the temperature coefficient of resistivity of tungsten after cold work increases with the annealing temperature from about 0.36 up to 0.49 per cent per C, at 2,500 C. Fig. 4 shows the relationship the authors have found between coefficient and annealing temperature for the tungsten wire as used in the thermometer. It has proved to be practical to anneal to a predetermined temperature and thus adjust the temperature coefficient to the required accuracy. The time and temperature of the anneal are initially chosen to provide adequate stability within the temperature rating of the thermometer, and the temperature coefficient thus established is maintained by suitable control of the anneal.

The resistance adjustment of resistance thermometers is usually performed mechanically. The wire is cut to give the proper value or has its cross section reduced by filing after the anneal. This adjustment process results in further strain, which must be removed by additional annealing before stability is assured. This can be a long and difficult process when high precision is desired. It has been found possible to adjust the tungsten bulb by a chemical process which can be carried out almost automatically. This adjustment follows the annealing previously described and is also done on the complete assembly, with the tungsten sensing element inside the stainless-steel tube. Extremely high adjustment accuracy is possible, and the advantages over the mechanical method are readily apparent.

At the completion of the coefficient and resistance adjustment, the bulbs are filled with a fine ceramic powder, which is

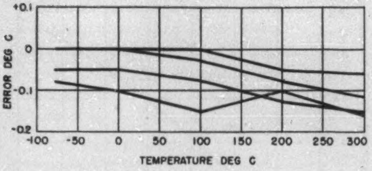


Fig. 5. Total calibration error in four tungsten resistance thermometers

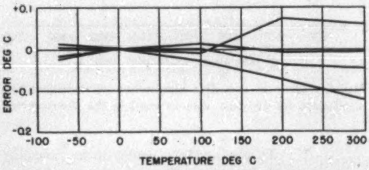


Fig. 6. Portion of calibration error in four tungsten thermometers caused only by variation in temperature coefficient

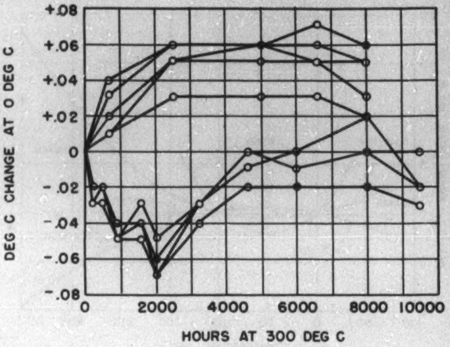


Fig. 8 (above). Stability tests showing variation of resistance at 0 C after exposure to 300 C

poured in through the seal tube. This powder is packed as tightly as possible by vibrating the tube during filling. The air remaining in the voids of the filling is then pumped out and replaced with pure helium gas, and the thermometer is ready for use.

Test Results

Fig. 5 shows the actual calibration errors of four experimental tungsten resistance thermometers manufactured as outlined above. Actual stated accuracies of these thermometers, if sold as commercial products, will of course exceed these figures. However, the results do indicate the accuracies that can be attained with moderate care. Calibration charts have been calculated over the −200 to 300 C range, from the familiar relationship $R_t = R_0(1+At+Bt^2+...)$. Maximum accuracy has been obtained by breaking up the calibration curves into two parts, one extending from −200 to 0 C and the other from 0 to 300 C. Suitable A and B constants were chosen for each of the two portions, to give maximum correlation with the observed results at specific temperatures, well scattered through the two ranges. The maximum deviation of the observed values from the calculated equation values was 0.005 C.

It is interesting to separate the calibration error of these bulbs into that amount due to incorrect temperature coefficient and that amount due to incorrect resistance adjustment. In order to do this, it is necessary to mathematically correct the resistance to the exact adjustment value, and observe the residual error which is due solely to an incorrect coefficient. Fig. 6 shows the results of this computation, and illustrates the effectiveness of the temperature-coefficient setting procedure.

The difference in the total calibration errors of Fig. 5 and the errors shown in Fig. 6 are caused by deviations in the resistance adjustment. Fig. 7 shows the errors that would be found at various temperatures for three different errors in resistance adjustment at the 0 C point. The errors in resistance adjustment have been converted to corresponding temperature errors.

In addition to accurate initial readings, the bulbs must maintain accuracies for long periods of time. In general, resistance thermometers show increasing instabilities at higher temperatures. For this reason, stability tests were conducted at the maximum rated temperature of the thermometer. Fig. 8 shows the change in calibration at the 0 C point after exposure to 300 C for the hours indicated. The errors shown include both changes in resistance and changes in the temperature coefficient. The changes in the temperature coefficient were measured separately at the end of the test shown in Fig. 8. A maximum change in resistance coefficient (per cent per C) of 0.0004, a minimum of 0.0000, and an average of 0.0002, were noted.

During the stability tests shown in Fig. 8, a standardized platinum bulb was used as a reference to check the test equipment and the 0 C reference temperature. It was interesting to observe that the best of the tungsten thermometers often showed greater stability between tests than the platinum standard. This was additionally impressive, since the platinum standard was maintained at room temperature, while the tungsten units were at 300 C!

The average response time of the bulbs, as defined by a 63-per-cent resistance change from 100 to 0 C, was 3.5 seconds. Smaller - diameter experimental bulbs, with a response time of less than 2 seconds have been built using the same construction principles.

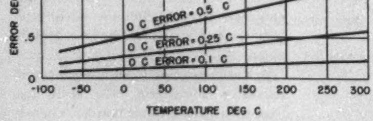


Fig. 7. Calibration errors as affected by accuracy of resistance adjustment at 0 C

Conclusions

Data has been presented to show the performance of a successful tungsten resistance thermometer which has been calibrated and adjusted by a unique method. In order to accomplish this it was necessary to develop new techniques, including a welding method which would produce low-resistance high-stability connections between the tungsten and nickel lead wires. More important, however, is the fact that tungsten has been successfully used in a commercial precision resistance thermometer for the first time. Although the present upper temperature rating is 300 C, this is not an inherent limitation in the sensing material, and progress is already being made in increasing this temperature.

The inherent temperature limitations of the commonly used metals, such as copper, platinum, nickel, etc., are well known. The best of these materials, platinum, has an upper range of approximately 600 C. However, experiments conducted with tungsten show excellent stability at temperatures up to 1000 C. Since the range is no longer limited by the inherent properties of the sensing element material, it can be expected that the upper temperature limit of resistance thermometers will be increased during the next few years. With the availability of better resistance thermometers, many new applications will be able to realize the better accuracy, better stability, and lower over-all cost of resistance thermometer equipment.

References

1. TUNGSTEN: A TREATISE ON ITS METALLURGY, PROPERTIES, AND APPLICATIONS (book), **Colin J. Smithells.** Chapman and Hall Ltd., London, England, 1952.

2. DIFFUSION OF GASES THROUGH SOLIDS, **Francis J. Norton.** *General Electric Review*, Schenectady, N. Y., vol. 55, Sept. 1952, pp. 28–29.

No Discussion

Fast-Response Magnetic Amplifiers

GERALD E. HUGHES
NONMEMBER AIEE

HARVEY A. MILLER
ASSOCIATE MEMBER AIEE

AMPLIFIERS for use in automatic control systems should feature high gain, high-power output, and fast response along with light weight and reliable operation. In electrical systems magnetic amplifiers are gaining favor because of their rugged construction, but until recently the response time of magnetic amplifiers has been relatively unsatisfactory for small servo systems. Amplifiers are often characterized by a time constant, but for carrier devices the number of cycles for 100 per cent response is sometimes a preferable characterization. In an elementary carrier type amplifier 100 per cent response can be obtained in 1/2 to 3/2 cycles depending on the time within the cycle at which the input change is applied. Gas tubes, for instance, achieve this result. This general behavior is often spoken of as single-cycle or half-cycle response. In this paper fast-response magnetic amplifiers are considered to be those which respond fully to an input change within 2 cycles of the carrier frequency.

Conventional self-saturating or external feedback amplifiers respond exponentially to a step change in input. If the attainment of 95 per cent of the ultimate value is considered full response, some of these amplifiers can be made to satisfy the above definition but only by sacrificing power gain. The load and control circuits are mutually coupled at all times, and this results in a relationship between power gain and time constant.

The elementary half-wave self-saturating magnetic amplifier inherently responds fully within 2 cycles since a reference condition is established each cycle. During load current conduction, the control circuit inductance approaches zero and the control current thus becomes dependent only on the control voltage and circuit resistance. In this amplifier, for a short period during each cycle, the control and load circuits are not mutually coupled, and this allows full response within, at most, 2 cycles. The degree of coupling affects the gain but not the response time. This mode of operation was recognized by Lufcy, et al,[1] and used in a fast-response magnetic servoamplifier.

A third fast-response magnetic amplifier was introduced by Ramey[2,3,4] in which the load and control circuits are not mutually coupled in the usual sense. Instead they operate on the magnetic state of the core on a time-sharing basis. This paper considers the time response and the operation of the basic Ramey circuit. The effects of control circuit resistance and winding limitations on the linearity, maximum realizable voltage gain, and power output are discussed in detail and compared with experimental

Nomenclature

a, b, c, d, e, f, g = subscripts indicating time sequence
A = load winding conductor cross section, circular mils (CM)
A_c = core cross section, inches^2
A_w = usable window area, inches^2
A_{wc} = control winding area inches^2
A_{wire} = control winding conductor cross section, CM
B_s = flux density at saturation
e_{ac} = supply voltage
e_o = output voltage
e_s = signal voltage
E_{ac} = peak supply voltage
E_o = peak output voltage
E_s = peak signal voltage
G = voltage gain
G_{max} = maximum realizable voltage gain
G_R = realizable gain
H = magnetizing force, ampere-turns per inch
H_c = coercive force, ampere-turns per inch at the power frequency
I_m = control circuit magnetizing current, amperes
I_G = load circuit current, amperes
J = current density, amperes per CM
k = peak volts per turn
K = fraction of window area for load winding
L_ϕ = mean length of flux path, inches
L_t = mean length of turn, feet
$n = (N_G/N_c)$ = turns ratio
N_c = control-winding turns
N_G = load-winding turns
$P_{dc\pi}$ = power output, direct current (conduction angle $= \pi$)
P_{dc} = power output, direct current (watts)
R_{cT} = total control circuit resistance
R_L = load resistance
R_w = load-winding resistance
t = time
T = period of supply voltage (sec)
v_c = control-winding voltage
v_G = load-winding voltage
X, see equation 24(A)
ρ = resistivity, ohm-CM per foot
$\Delta\phi$ = increment of flux (maxwells)
ϕ = flux (maxwells)
ϕ_s = saturation flux (maxwells)
ω = angular frequency of supply voltage

Paper 54-159, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 8, 1953; made available for printing December 15, 1953.

GERALD E. HUGHES and HARVEY A. MILLER are with the Raytheon Manufacturing Company, Waltham, Mass.

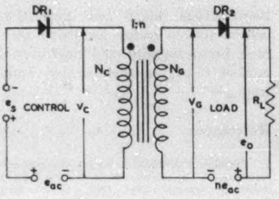


Fig. 1. Ideal circuit

data. The maximum gain and power output are expressed in terms of core dimensions and material constants.

Idealized Circuit Operation and Gain

The relations describing the operation of the ideal circuit are derived to facilitate the later treatment of the more difficult nonideal circuit. The operation is described under the following assumptions:

1. The core displays a rectangular flux-current loop of negligible width.
2. Ideal rectifiers are used.
3. The voltage sources are sinusoidal.
4. The winding resistances are zero.
5. Steady-state operation exists.

The basic circuit is shown in Fig. 1 and the assumed flux-current loop in Fig. 2. The voltages e_{ac} and ne_{ac} are sinusoidal, while e_s is a half-wave rectified sinusoid. In this paper e_s is considered to be the signal voltage. The ideal rectifiers are designated DR_1 and DR_2.

The initial flux is chosen at point a in Fig. 2 since this point is common to all cycles during operation of the circuit. The control and load functions can be considered individually, since, during the entire positive half-cycle of supply voltage, the rectifier DR_2 is blocked and hence no voltage or current is supplied to the load. In the control circuit, however, a voltage $(e_{ac}-e_s)$ appears across the control winding. The resulting flux change can be obtained from Faraday's law

$$\left.\begin{aligned} v_c &= e_{ac}-e_s & (1)\\ \Delta\phi_{ad} &= -\frac{10^8}{N_c}\int_{t_a}^{t_d} v_c dt & (2)\end{aligned}\right\} 0 \leq |e_s| \leq |e_{ac}|$$

The voltage magnitude restriction is satisfied in conventional circuit operation.

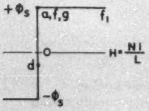


Fig. 2. Ideal flux-current loop

At time t_d as shown in Fig. 3, the positive half-cycle ends and the flux remains at some point d (see Fig. 2) determined by equations 1 and 2. The flux is not zero since the core material is capable of acting as a permanent magnet. This entire process is known as resetting the flux, and this half-cycle is called the reset time.

During the negative half-cycle of supply voltage, the rectifier DR_1 is blocked, but the load circuit is operative. With reference to Figs. 1 and 3, the following equations hold during the negative half-cycle.

$$\left.\begin{aligned} e_o &= -ne_{ac}+v_G = 0 & (3)\\ \Delta\phi_{df} &= -\frac{10^8}{N_G}\int_{t_d}^{t_f} v_G dt & (4)\end{aligned}\right\} t_d<t<t_f$$

The core saturates at time t_f determined by

$$\Delta\phi_{ad}+\Delta\phi_{df}=0 \qquad (5)$$

and until the end of the half-cycle

$$e_o=-ne_{ac}, \qquad t_f<t<t_g \qquad (6)$$

By substituting

$e_{ac}=E_{ac}\sin\omega t$
$e_s=E_s\sin\omega t$
$t_a=0$
$t_d=\pi/\omega$ seconds
$t_g=2\pi/\omega$ seconds

Equations 1 through 6 can be solved for e_o in terms of e_s and e_{ac}. Equations 1 and 2 give the amount of flux resetting.

$$\Delta\phi_{ad}=-\frac{10^8}{N_c}\int_0^{\pi/\omega}(E_{ac}-E_s)\sin\omega t dt = -\frac{10^8}{N_c\omega}2(E_{ac}-E_s) \qquad (7)$$

Using equation 5 the flux change during the load half cycle is determined

$$\Delta\phi_{df}=\frac{10^8}{N_c\omega}2(E_{ac}-E_s) \qquad (8)$$

If equation 3 is integrated, equation 4 may be substituted for the integral of v_G and an expression involving the output voltage and $\Delta\phi_{df}$ is obtained.

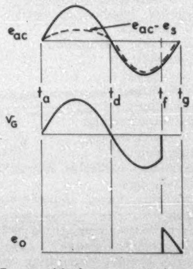


Fig. 3. Ideal circuit wave forms

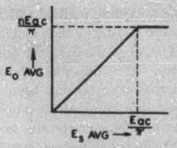


Fig. 4. Ideal transfer curve

$$\int_{\pi/\omega}^{2\pi/\omega} e_o dt = -\int_{\pi/\omega}^{2\pi/\omega} nE_{ac}\sin\omega t dt + \int_{\pi/\omega}^{t_f} v_G dt - \int_{t_f}^{2\pi/\omega}(0)dt = \frac{2nE_{ac}}{\omega} - \frac{N_G\Delta\phi_{df}}{10^8} \qquad (9)$$

Equations 8 and 9 then yield

$$\int_{\pi/\omega}^{2\pi/\omega} e_o dt = \frac{+2nE_{ac}}{\omega} - \frac{2nE_{ac}}{\omega} + \frac{2nE_s}{\omega} = \frac{2nE_s}{\omega} \qquad (10)$$

The average output voltage is given by

$$E_{o\,av}=\frac{1}{T}\int_0^{2\pi/\omega} e_o dt = \frac{\omega}{2\pi}\int_{\pi/\omega}^{2\pi/\omega} e_o dt = \frac{nE_s}{\pi} \qquad (11)$$

Since the average value of the half-wave rectified signal voltage is E_s/π, equation 11 may be expressed as

$$E_{o\,av}=+nE_{s\,av} \qquad (12)$$

which describes the output-input or transfer curve of the amplifier.

The voltage gain will be defined as the slope of the transfer curve.

$$G=\left|\frac{dE_{o\,av}}{dE_{s\,av}}\right| = n = \frac{N_G}{N_c} \qquad (13)$$

The voltage gain is thus seen to be directly proportional to the turns ratio. Under ideal conditions it would be desirable to make N_G as large and N_c as small as possible. The transfer and gain curves expressed by equations 12 and 13 are shown in Figs. 4 and 5.

Under the assumptions made, there is no control current and hence no power input; therefore, power, current, and ampere-turn gain are meaningless. Note that the derivation required that the entire voltage $(e_{ac}-e_s)$ appear across the control winding. This can be accomplished only by using a perfect voltage source and a zero resistance circuit, or by using an imperfect source supplying no current.

It is the time integral of e_s that controls the average output voltage. The time for 100 per cent response is maxi-

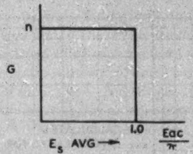


Fig. 5. Ideal gain curve

mized at 2 cycles for a change in e_s at $t=t_a+$ and minimized at 1/2 cycle for a change in e_s at $t=t_a-$. For circuits using more than one basic amplifier, this response time can be reduced.

Gain Under Nonideal Conditions

CONTROL CIRCUIT RESISTANCE AND CORE LOSS EFFECTS

The previous gain derivation involved a lossless core and a perfect voltage source. Neither of these conditions is likely to be met in practice. Commercially available core materials such as Deltamax, Hypernik V, Orthonik, Orthonol, Permeron, etc., exhibit flux-current loops with extremely steep sides and a width which is nearly constant[5,6,7] at power frequencies commonly used in control equipment. The area of the loop is indicative of the loss per cycle. The assumption of perfect rectifiers is more nearly met, but it is required that the maximum reverse current of the rectifier be less than half the width of the flux-current loop to avoid spurious resetting action.

The purpose of this section is to determine the effect of core loss and control source resistance on the transfer and gain curves with particular emphasis on the gain and its limitations.

The assumptions made are:

1. The core displays a rectangular flux-current loop with constant width for partial flux excursions.
2. Ideal rectifiers are used.
3. The voltage sources are sinusoidal.
4. Steady-state operation exists.

The assumed flux-current loop is shown in Fig. 6 and the circuit in Fig. 7. All control circuit resistance is lumped into R_{cT} and assumed constant. This includes source resistance, winding resistance, and rectifier forward resistance.

The circuit equations for Fig. 7 for the positive or reset half cycle are

$$(E_{ac}-E_s)\sin\omega t=v_c+I_mR_{cT} \tag{14}$$

and

$$\Delta\phi_{ad}=-\frac{10^8}{N_c}\int_{t_a}^{t_d}v_c\,dt \tag{15}$$

The magnetizing current I_m is defined by

$$N_cI_m=H_cL_\phi \tag{16}$$

where L_ϕ is the mean length of the magnetic flux path.

The flux positions and wave forms are shown in Figs. 6 and 8. At time t_a, the beginning of the reset half-cycle, the current I_m will rise, limited only by resistance R_{cT}, until time t_b when the side of the flux-current loop is reached. At this time

$$\frac{E_{ac}-E_s}{R_{cT}}\sin\omega t_b=I_m=\frac{H_cL_\phi}{N_c} \tag{17}$$

The control current will remain at this value until time t_c near the end of the half-cycle since the incremental inductance is extremely large. The current then decreases to zero with the applied voltage. During the time interval $t_b<t<t_c$, a voltage will appear across the control winding and the core flux will change according to equations 14 and 15. From equation 17

$$t_b=\frac{1}{\omega}\sin^{-1}\frac{I_mR_{cT}}{E_{ac}-E_s}\text{ in quadrant 1} \tag{17A}$$

Similarly

$$t_c=\frac{1}{\omega}\sin^{-1}\frac{I_mR_{cT}}{E_{ac}-E_s}=\frac{\pi}{\omega}-t_b\text{ in quadrant 2} \tag{17B}$$

The manipulation to find the flux change during reset from equation 15 is carried out in Appendix I and results in

$$\Delta\phi_{ad}=-\frac{2}{\omega N_c}10^8(E_{ac}-E_s)\times\left[\cos\omega t_b-\left(\frac{\pi}{2}-\omega t_b\right)\sin\omega t_b\right] \tag{18}$$

At time t_d the half-cycle ends and the flux remains at some point determined by equation 18.

When the negative half-cycle is applied, the current I_G rises, limited only by the resistance in the load circuit, until time t_e when

$$\frac{nE_{ac}\sin\omega t_e}{R_L}=I_G=\frac{H_cL_\phi}{N_G} \tag{19}$$

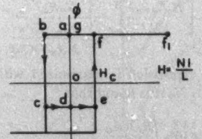


Fig. 6. Assumed flux-current loop

Lower case letters correspond to time subscripts in Fig. 8

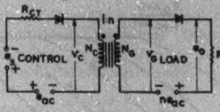


Fig. 7 Nonideal circuit

At this time, voltage begins to appear across the load winding N_G and the following equations hold

$$ne_{ac}=v_G-e_o \tag{20}$$

$$\Delta\phi_{df}=-\frac{10^8}{N_G}\int_{t_e}^{t_f}v_G\,dt \tag{21}$$

At $t=t_f$, saturation occurs since

$$\Delta\phi_{ad}+\Delta\phi_{df}=0 \tag{22}$$

and thereafter

$$e_o=-ne_{ac} \tag{23}$$

At the end of the half-cycle core flux is again at point a in Fig. 6.

Solution of these equations for $E_{0\text{ av}}$ is carried out in Appendix II and the result is

$$\frac{E_{0\text{ av}}}{E_{ac\text{ av}}}=n\left[1-(1-E_s/E_{ac})\left\{\sqrt{1-X^2}-\left(\frac{\pi}{2}-\sin^{-1}X\right)X\right\}\right] \tag{24}$$

where

$$X=\frac{H_cL_\phi R_{cT}}{kN_c^2(1-E_s/E_{ac})}=\sin\omega t_b \tag{24A}$$

The per unit transfer curves defined by equation 24 are plotted in Fig. 9 for different R_{cT}. The loss in linearity and shift in operating point as R_{cT} is increased are readily seen, while the change in slope is not as apparent.

The gain of the circuit under these

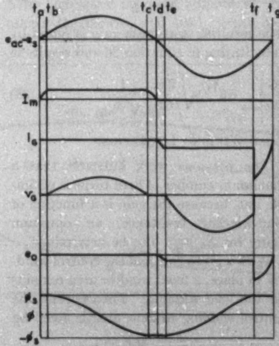


Fig. 8. Nonideal circuit wave forms

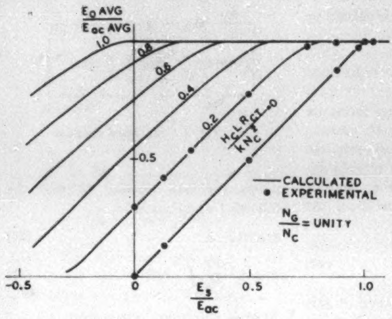


Fig. 9. Output-input transfer curves

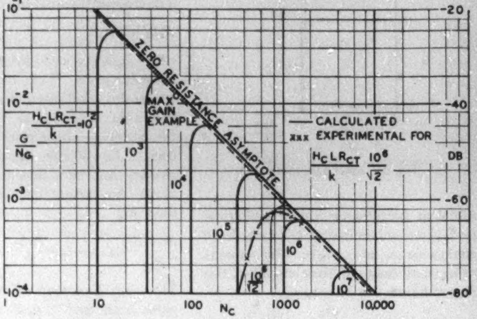


Fig. 10. Gain versus control turns with $R_{cT}[H_cL/k]$ as a parameter

conditions is again $(dE_{o\ av})/(dE_{s\ av})$ and is obtained from equation 24 in Appendix III. The gain is

$$G=n\sqrt{1-X^2}=\frac{N_G}{N_c}\times \sqrt{1-\left[\frac{H_cL_\phi R_{cT}}{kN_c^2(1-E_s/E_{ac})}\right]^2} \qquad (25)$$

H_cL_ϕ/k is a function of the core only and R_{cT} is a constant depending on the choice of control source. Equation 26 illustrates how gain per load turn varies with N_c for zero signal and different values of the constant $H_cL_\phi R_{cT}/k$ and is shown in Fig. 10.

$$\frac{G}{N_G}=\frac{1}{N_c}\sqrt{1-\left[\frac{H_cL_\phi R_{cT}}{kN_c^2}\right]^2} \qquad (26)$$

It is evident that with a fixed core, a fixed R_{cT}, and the zero signal condition, there exists a maximum gain to be obtained by varying N_c. As shown in Appendix IV, the required number of control turns for maximum gain is

$$(N_c)_{\text{max gain}}=3^{1/4}\sqrt{\frac{H_cL_\phi R_{cT}}{k}} \qquad (27)$$

The resulting maximum gain is found by substituting in equation 26 and results in

$$\left.\frac{G}{N_G}\right|_{\max}=3^{-3/4}\sqrt{\frac{2k}{H_cL_\phi}}\,\frac{1}{\sqrt{R_{cT}}} \qquad (28)$$

LOAD-WINDING LIMITATIONS

The previous work indicates that a maximum number of load turns is desirable but, because the gain is a function of load-winding resistance, an optimum figure for N_G may also be determined.

With a given core and the control winding in place, a fixed window area remains for the load winding. Theoretically the control winding window area is given by

$$A_{wc}=N_cA_{\text{wire}}\frac{\pi}{4\times 10^6}=\frac{N_cI_m}{J}\,\frac{\pi}{4\times 10^6} \qquad (29)$$

Since N_cI_m is a constant of the core used, A_{wc} is a constant for a given J. The called-for wire size is often much too small for practical use so that the control winding area is greater than indicated above. In cases of low R_{cT} the control winding area is often a negligible part of the available window area.

The resistance of the load winding is

$$R_w=\rho\frac{N_GL_t}{A} \qquad (30)$$

while the winding area required must be equal to or less than that available

$$N_G\frac{A\pi}{4\times 10^6}\leqq KA_w \qquad (31)$$

Thus the load-winding resistance is proportional to N_G^2 as shown in equation 32

$$R_w\geqq\rho\frac{L_t\pi}{4KA_w10^6}N_G^2 \qquad (32)$$

The inequality indicates that if the window area is not filled R_w may be greater than the minimum obtainable for that value of N_G.

The realized voltage gain is $R_L/(R_w+R_L)$ of the theoretical gain since R_w has not been taken into account. Thus from equation 26 the realized gain becomes

$$G_R=\left|\frac{G}{N_G}\right|N_G\frac{R_L}{R_w+R_L}\leqq\frac{1}{N_c}\times \sqrt{1-\left[\frac{H_cL_\phi R_{cT}}{kN_c^2}\right]^2}\,\frac{N_GR_L}{R_L+\dfrac{\rho L_tN_G^2\pi}{KA_w4\times 10^6}} \qquad (33)$$

If this expression is maximized with respect to N_G it is found that

$$N_G=2\times 10^3\sqrt{\frac{R_LKA_w}{\rho L_t\pi}} \qquad (34)$$

for maximum voltage gain by load circuit design. This also means $R_w=R_L$ from equation 32. The over-all maximum realizable voltage gain becomes from equations 28, 33, and 34

$$G_{\max}=3^{-3/4}10^3\sqrt{\frac{2k}{H_cL_\phi}\,\frac{KA_w}{\rho L_t\pi}\,\frac{R_L}{R_{cT}}} \qquad (35)$$

Equation 27 shows the way in which to achieve maximum gain through control circuit design, while equation 34 again maximizes gain for load circuit design. If the validity of the assumptions is kept in mind, equation 35 shows the maximum gain to be expected.

EXAMPLE

To illustrate the method for obtaining maximum voltage gain, an available toroidal core will be chosen for an example (Arnold Engineering Company Deltamax no. *4180*). The constants are approximately

$L_\phi=8.65$ inches
$H_c=0.674\ \dfrac{\text{ampere-turns}}{\text{inch}}$
$A_c=0.125$ inch²
$B_s=90.3\ \dfrac{\text{kilolines}}{\text{inches}^2}$
$k=0.0425\ \dfrac{\text{peak volts}}{\text{turn}}$ at 60 cycles per second
$A_w=2.84$ inches²

The actual window area for this core is 4.37 inches,² but only 2.84 inches² is usable allowing for winding machine requirements. This allowance must be determined by experience. With an assumed total source resistance of 20 ohms, equations 27 and 28 lead to

$N_c=69$ turns

and

$$\left|\frac{G}{N_G}\right|_{\max}=1.18\times 10^{-2}\text{ per load turn}$$

Assuming R_L equal to 100 ohms a maximum N_G is desired within the available window area and with $R_w=R_L$. The

69-turn control-winding will take very little area and, if the usable window area is completely filled with the load winding (no insulation allowance) and if a conservative 4-inch mean length per turn is chosen, the G_{max} from equation 35 becomes 60.2. From equation 34, N_G is 10,200.

Using a standard wire table which allows for wire insulation[8] and random winding practice, N_G may be approximately 6,900 turns of no. 25 enameled wire with $R_w = 74.5$ ohms, or 8,750 turns of no. 26 enameled wire with $R_w = 119$ ohms. Choosing the latter, the supply voltage ne_{ac} is 275 volts rms and the realized gain is by equation 33

$$G_R = 47.2$$

It is 46.5 for the larger wire. This gives the approximate voltage gain to be expected using the assumed circuit and components.

Power Output

Where the magnetic amplifier is to deliver appreciable average power, the current-carrying capacity of the wire in the load-winding becomes important. The use of realistic figures for the current density J in the load winding will require that the output of the maximum voltage gain amplifier be limited to a small fraction of its maximum possible voltage output, i.e., the conduction angle must be much less than π radians. This results in an even smaller fraction of the maximum possible power output into a given load. It will be shown that a small sacrifice in voltage gain can result in a considerable increase in power output.

The useful power output is assumed to be due to the direct-current component. This assumption may not be justified in many cases. For complete half-cycle conduction into the load

$$P_{dc\,\pi} = I_{av}{}^2 R_L = \left[\frac{nE_{ac}}{\pi(R_L + R_w)}\right]^2 R_L \qquad (36)$$

By substituting for nE_{ac} and for R_w from equation 32 $P_{dc\,\pi}$ may be expressed in terms of the core constants and N_G.

$$P_{dc\,\pi} = \left[\frac{kN_G}{\pi\left(R_L + \frac{\rho L_t \pi}{KA_w 4\times 10^6} N_G{}^2\right)}\right]^2 R_L \qquad (37)$$

In obtaining this power output, the current-carrying capacity of the load winding N_G must be considered. If the allowable current density is J, the maximum current is

$$I_{av\ max} = JA \qquad (38)$$

In selecting J consideration must be given to the heating effect of harmonic currents. From equation 31 there can be $N_G = KA_w 4\times 10^6/A\pi$ turns in a given window area KA_w. Equation 38 becomes

$$I_{av\ max} \leqq \frac{JKA_w 4\times 10^6}{N_G \pi} \qquad (39)$$

which determines an upper limit for the power output. Thus

$$P_{dc\,\pi} = \left[\frac{kN_G}{\pi\left(R_L + \frac{\rho L_t \pi}{KA_w 4\times 10^6} N_G{}^2\right)}\right]^2 R_L = \left[\frac{4JKA_w 10^6}{N_G \pi}\right]^2 R_L \qquad (40)$$

This may be expressed in per unit values

$$(\mathrm{pu}P_{dc\,\pi}) = \frac{4(\mathrm{pu}N_G)^2}{\{1 + (\mathrm{pu}N_G)^2\}^2} = 4\left[\frac{\pi J \rho L_t}{k(\mathrm{pu}N_G)}\right]^2 \qquad (41)$$

Unit N_G is given by equation 34 and unit $P_{dc\,\pi}$ is found by substituting unit N_G in equation 37

$$\text{unit } P_{dc\,\pi} \equiv \frac{k^2 KA_w 10^6}{\pi^3 \rho L_t} \qquad (42)$$

The per unit power output is shown in Fig. 11 as a function of $(\mathrm{pu}N_G)^2$ along with the allowable power output due to current density limitations. Solution of equation 41 for N_G will locate the intersection of the two curves and is

$$(\mathrm{pu}N_G)^2 = \frac{\pi J \rho L_t/k}{1 - \pi J \rho L_t/k} \qquad (43)$$

The solid curve indicates the per unit voltage gain where unit gain is given by equation 35. The curves show that, if N_G is halved and the wire size increased to fill the window area, a 20-per-cent loss in voltage gain is suffered but that the maximum allowable power output has been increased for common values of the parameter $k/J\pi\rho L_t$.

For the example previously considered ($N_G = 10{,}200$) the per unit gain is 1 (corresponding to 60.2) but the allowable per unit power with $J = 10^{-3}$ is only 0.356 (17 watts).

If equation 43 is satisfied

$$(\mathrm{pu}N_G)^2 = \frac{\frac{1}{3.91}}{1 - \frac{1}{3.91}} = 0.355 \qquad (44)$$

$$N_G = 10{,}200\sqrt{0.355} = 6{,}080$$

the gain becomes 0.88 while the power output becomes 0.78 (37.2 watts). This is a desirable compromise in many cases.

With American Wire Gauge wire sizes and including allowance for random winding, the foregoing figures cannot be realized. Fig. 12 shows the gain and power output for this case. No. *23* wire gives maximum power output of 0.430 (20.5 watts) while the voltage gain is 0.657 (39.5). The decrease is due to the insulation and commercial practice allowances.

Remarks

The circuit considered is one of the basic high-speed magnetic amplifiers. It allows the control and load circuit de-

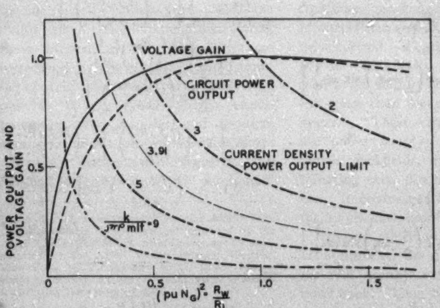


Fig. 11 Theoretical voltage gain and power output

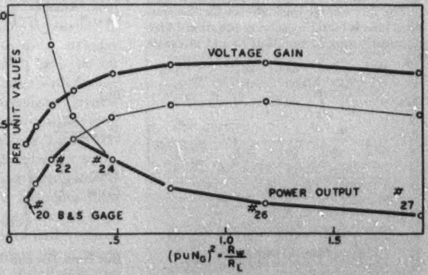


Fig. 12. Voltage gain and power output for a random wound toroid using American Wire Gauge wire

signs to be considered separately since they operate on the core on a time-sharing basis.

Maximum voltage gain by control circuit design is obtained by selecting a rectangular flux-current-loop core material with a minimum coercive force and a maximum of flux-carrying capacity. The flux may be increased by increasing the width of the toroid but this reduces the effective rectangularity of the flux-current loop.[6,7] An increase in height results in a larger mean length of turn. With a given core a method of designing for maximum voltage gain is presented. This gain is proportional to $1/\sqrt{R_{cT}}$.

Load circuit design is dependent on the winding area available and because of this restriction, maximum voltage gain is achieved by equating R_w to R_L. The maximum gain is proportional to $\sqrt{R_L}$. Gains of the order of 45 are practicable with R_L of 100 ohms and with a control circuit resistance of 20 ohms.

In Fig. 10, an experimental curve of gain per load turn is plotted for the core used in the example and for $H_cL\,R_{cT}/k = 10^8/\sqrt{2}$. The comparison shows some discrepancy between the theoretical and experimental maximum gain but the number of control turns required to obtain maximum gain compares favorably considering the departures of the physical circuit from the theoretical assumptions.

The maximum power output is usually limited by the allowable current density in the load winding. If a small sacrifice is made in voltage gain, the maximum power output may be greatly increased. In the example presented, a theoretical increase in maximum power output from 17 to 37.2 watts is obtained by accepting a voltage gain 88 per cent of the maximum of 60.2.

Appendix I

The flux change during reset for the nonideal case is found by solving equation 14 for v_c, substituting in equation 15 and integrating

$$\Delta\phi_{ad} = -\frac{10^8}{N_c}\left[\int_0^{\pi/\omega}(E_{ac}-E_s)\sin\omega t\,dt - \int_0^{\pi/\omega} I_mR_{cT}\,dt\right]$$

$$= -\frac{10^8}{N_c}\left[-\frac{(E_{ac}-E_s)}{\omega}\cos\omega t\Big|_0^{\pi/\omega} - 2\int_0^{t_b}(E_{ac}-E_s)\sin\omega t\,dt - \int_{t_b}^{t_c} I_mR_{cT}\,dt\right]$$

$$\Delta\phi_{ad} = -\frac{10^8}{N_c}\left[\frac{2}{\omega}(E_{ac}-E_s)+2\frac{E_{ac}-E_s}{\omega}\times(\cos\omega t_b-1)-I_mR_{cT}(t_c-t_b)\right]$$

But

$$t_c = \frac{\pi}{\omega} - t_b$$

and from equation 17

$$I_mR_{cT} = (E_{ac}-E_s)\sin\omega t_b$$

Thus

$$\Delta\phi_{ad} = -\frac{10^8}{N_c}\left[\frac{2}{\omega}(E_{ac}-E_s)+2\frac{E_{ac}-E_s}{\omega}\times(\cos\omega t_b-1)-(E_{ac}-E_s)\left(\frac{\pi}{\omega}-2t_b\right)\times\sin\omega t_b\right]$$

$$= -\frac{2}{\omega N_c}10^8(E_{ac}-E_s)\left\{1+\cos\omega t_b-1-\left(\frac{\pi}{2}-\omega t_b\right)\sin\omega t_b\right\} \quad (18)$$

$$\Delta\phi_{ad} = -\frac{2}{\omega N_c}10^8(E_{ac}-E_s)\times\left\{\cos\omega t_b-\left(\frac{\pi}{2}-\omega t_b\right)\sin\omega t_b\right\}$$

Appendix II

From equation 20

$$E_{o\,av} = \frac{\omega}{2\pi}\int_{\pi/\omega}^{2\pi/\omega} e_o\,dt = \frac{\omega}{2\pi}\int_{\pi/\omega}^{2\pi/\omega}\times(v_G-ne_{ac})\,dt$$

Substituting from equation 21

$$E_{o\,av} = \frac{\omega}{2\pi}\left[-\frac{\Delta\phi_{df}N_G}{10^8}+\frac{2nE_{ac}}{\omega}\right]$$

From equations 22 and 18

$$E_{o\,av} = \frac{\omega}{2\pi}\left[-N_G\frac{2}{\omega N_c}(E_{ac}-E_s)\times\left\{\cos\omega t_b-\left(\frac{\pi}{2}-\omega t_b\right)\sin\omega t_b\right\}+\frac{2nE_{ac}}{\omega}\right]$$

$$E_{o\,av} = \frac{nE_{ac}}{\pi}\left[1-\left(1-\frac{E_s}{E_{ac}}\right)\times\left\{\cos\omega t_b-\left(\frac{\pi}{2}-\omega t_b\right)\sin\omega t_b\right\}\right]$$

But

$$E_{ac\,av} = \frac{E_{ac}}{\pi}$$

$$\frac{E_{o\,av}}{E_{ac\,av}} = n\left[1-\left(1-\frac{E_s}{E_{ac}}\right)\times\left\{\cos\omega t_b-\left(\frac{\pi}{2}-\omega t_b\right)\sin\omega t_b\right\}\right]$$

But from 17*a* and 17

$$\sin\omega t_b = \frac{I_mR_{cT}}{E_{ac}-E_s} = \frac{H_cL_\phi R_{cT}}{N_c(E_{ac}-E_s)}$$

For *k* volts per turn

$$\sin\omega t_b = \frac{I_mR_{cT}}{E_{ac}-E_s} = \frac{H_cLR_{cT}}{kN_c^2\left(1-\frac{E_s}{E_{ac}}\right)} \equiv X$$

and

$$\cos\omega t_b = \sqrt{1-X^2}$$

Thus

$$\frac{E_{o\,av}}{E_{ac\,av}} = n\left[1-\left(1-\frac{E_s}{E_{ac}}\right)\times\left\{\sqrt{1-X^2}-\left(\frac{\pi}{2}-\sin^{-1}X\right)X\right\}\right] \quad (24)$$

Appendix III

The voltage gain is found by differentiating equation 24.

$$G = \frac{dE_{o\,av}}{dE_{s\,av}} = \frac{dE_{o\,av}}{d\frac{E_s}{\pi}} = n\left[\sqrt{1-X^2}-X\left(\frac{\pi}{2}-\sin^{-1}X\right)-(E_{ac}-E_s)\times\left(\frac{-2X}{2\sqrt{1-X^2}}\frac{dX}{dE_s}-\left\{\frac{\pi}{2}-\sin^{-1}X\right\}\frac{dX}{dE_s}\right)-X\left(\frac{-X}{\sqrt{1-X^2}}\frac{dX}{dE_s}\right)\right]$$

But

$$\frac{dX}{dE_s} = \frac{I_mR_{cT}}{(E_{ac}-E_s)^2} = \frac{X}{E_{ac}-E_s}$$

Substituting into the expression for G, the result is

$$G = n\sqrt{1-X^2} = n\sqrt{1-\left[\frac{H_cL_\phi R_{cT}}{kN_c^2\left(1-\frac{E_s}{E_{ac}}\right)}\right]^2} \quad (25)$$

Appendix IV

The maximum gain condition is found by maximizing equation 26.

$$\frac{1}{N_G}\frac{dG}{dN_c} = -\frac{1}{N_c^2}\left[1-\left(\frac{H_cL_\phi R_{cT}}{kN_c^2}\right)^2\right]^{1/2} + \frac{1}{2N_c}\frac{\left(\frac{H_cL_\phi R_{cT}}{k}\right)^2\frac{4}{N_c^5}}{\left[1-\left(\frac{H_cL_\phi R_{cT}}{kN_c^2}\right)^2\right]^{1/2}} = 0$$

$$N_c^4\left[1-\left(\frac{H_cL_\phi R_{cT}}{kN_c^2}\right)^2\right] = 2\left(\frac{H_cL_\phi R_{cT}}{k}\right)^2$$

$$N_c^4 - N_c^4\left(\frac{H_cL_\phi R_{cT}}{kN_c^2}\right)^2 = 2\left(\frac{H_cL_\phi R_{cT}}{k}\right)^2$$

$$N_c^4 = 3\left(\frac{H_cL_\phi R_{cT}}{k}\right)^2$$

Thus

$$(N_c)_{\text{max gain}} = 3^{1/4}\sqrt{\frac{H_cLR_{cT}}{k}} \quad (27)$$

References

1. An Improved Magnetic Servo Amplifier, C. W. Lufcy, A. E. Schmid, P. W. Barnhart. *AIEE Transactions*, vol. 71, pt. I, Sept. 1952, pp. 281–89.

2. On the Mechanics of Magnetic Amplifier Operation, Robert A. Ramey. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1214–23.

3. On the Control of Magnetic Amplifiers, R. A. Ramey. *Ibid.*, pp. 2124–48.

4. High Speed Magnetic Amplifier, L. J. Johnson. *Electrical Manufacturing*, New York, N. Y., vol. 50, no. 5, Nov. 1952, p. 98.

5. The Effect of Core Materials on Magnetic Amplifier Circuits, Leo J. Johnson. *AIEE Transactions*, vol. 71, pt. I, 1952, pp. 26–31.

6. Evaluation of Core Materials for Magnetic Amplifiers, D. C. Dieterly. *Electrical Manufacturing*, New York, N. Y., vol. 51, no. 1, Jan. 1953, p. 68; no. 2, Feb. 1953, p. 124.

7. Dynamic Hysteresis Loops of Several Core Materials Employed in Magnetic Amplifiers, Harold W. Lord. *AIEE Transactions*, vol. 72, pt. I, March 1953, pp. 85–88.

8. Radio Engineers Handbook (book), F. E. Terman. McGraw-Hill Book Company, Inc., New York, N. Y., 1943, p. 103.

No Discussion

Mechanical Aspects and Component Features of a New 12-Channel Open-Wire Carrier System

A. G. EWING
NONMEMBER AIEE

F. W. FRAZEE
NONMEMBER AIEE

DALE WELLING
NONMEMBER AIEE

A NEW line of carrier equipment known as the class *45* has been developed with a view toward application on open-wire lines, paired and coaxial cable, and radio systems. Both mechanical and electrical design have been co-ordinated for each of the systems in these respective applications. In this paper, the *45A* system arrangements for open-wire use are described. In basic terminal equipment, no differences exist for the various applications. Substitution of a small number of common group-modulating assemblies, modifications in slight degree in carrier supply, and rearrangements of terminal frame assemblies identify specific uses. Great stress has been placed on standardization in this respect, both to speed the development in the different areas of ultimate use and also to increase the demand for specific items of manufacture. With the major part of this equipment alike among these different systems, and differences introduced by the relatively small number of group-identifying parts, it has been possible to come out almost simultaneously with manufactured terminal arrangements for most of the fields of application. Transmission features and electric performance of the equipment are described in a companion paper by Appert, Caruthers, and Chaskin.[1]

The equipment arrangements are miniaturized to a very high degree and extensive use is made of plug-in unit arrangements. To attain a maximum of uniformity, these various plug-in units make use of a number of plug-in subassemblies to provide needed differences.

All of the applications of *45*-class terminal equipment have common basic problems in signaling circuit use, need for automatic channel regulation, optional use of compandors, and basic channelizing circuits. This has made it simple to incorporate a design standardized for such uses, and it affords a simple, quick, and inexpensive method of applying options.

In addition to co-ordinating the system design for ultimate application, accessory, equipment associated with terminals and repeaters has also been modernized and miniaturized. This makes it possible to provide completely racked-up arrangements on factory-shipped bays in far less space than was required with former systems. These arrangements include a-c power supply, ringdown converter panels, jackfields, miniature 4-wire terminating sets, power distribution units, etc. In contrast to former arrangements for 12-channel terminals including these accessories which occupied several bays of equipment, the miniaturized versions for *45A* occupy a fraction of a single relay rack. Thus both installation and portability are far simpler than previously.

Terminal Equipment Arrangements

In Fig. 1 is shown the 12-channel terminal of a *45A* system. This is a unitized arrangement complete with terminal block, fuses, and alarms ready for mounting on a standard 19-inch telephone-type relay rack of channel or duct type. The system comes complete from the factory with field installation, requiring only the connection of speech and signal drop leads, high-frequency line, and power supply to the terminal block. Only 31$^1/_2$ inches of vertical rack space is required, and as many as four such terminals can be mounted in an 11-foot 6-inch relay rack. Full 10-inch depth of mounting is used with equipment projecting 5 inches to the front of the relay rack. This permits back-to-back arrangements of two *45A* bays with a total of 96 channels of equipment in the relay rack space between two aisles.

The shelf assembly shown consists of five sections or rows of equipment, largely

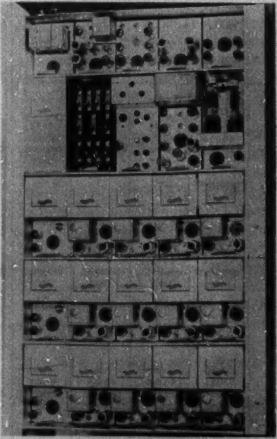


Fig. 1. Twelve-channel type-45A terminal shelf assembly

Paper 54-102, recommended by the AIEE Wire Communications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 22, 1953; made available for printing December 10, 1953.

A. G. Ewing, F. W. Frazee, and Dale Welling are with the Lenkurt Electric Company, Inc., San Carlos, Calif.

of a plug-in type. The one unit on the top section not plug-in is the carrier distribution unit at the upper left. In the second section, the terminal block at the left adjacent fuse panel, and next adjacent alarm panel all come as part of the wired shelf assembly. Each of these wired-in units, however, are connected into the wiring harness of the shelf at terminal board connections to the rear. Except for the terminal block, these units slide from the shelf for rear accessibility through use of hinged cables.

In Fig. 2, the arrangement of plug-in and wired units is shown. The bottom three sections of plug-in units are identical, each consisting of four channel units and one pregroup or combining unit for the associated four channels. Each plug-in unit in these three sections is a rectangular box assembly $3^1/_4$ inches wide, $6^3/_8$ inches high, and 10 inches in depth. Each has a 20-pin multicontact plug mounted at its rear through which all transmission and power connections are carried. Each unit plugs into a jack at the rear of the shelf assembly. Further detail on these particular plug-in units is shown in Fig. 1. A guide track arrangement allows each of the plug-in units to slide in a groove to make plug-and-jack connections. A guide pin on the jack and a shoe on the plug further align the contacts of the plug and jack. It is also necessary to keep the jack floating freely in its mounting arrangement. Each unit is cammed into a complete seating of the plug and disengaged without excessive pull, through use of a latch-type lever and gear on the bottom plate near the front of the unit.

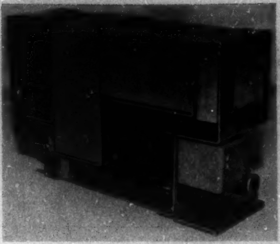


Fig. 3. Channel unit with all plug-in subassemblies in place

On the front of the plug-in channel units are square plug-in resistance hybrids. A gain control for setting 2-wire circuit equivalent is beneath the circular button on each hybrid. Alternative plug-in units are provided for 4-wire operation. Two types of plug-in units may be used, containing pads for operation in offices with −13, +4 levels, or −16, +7 levels. Removal of the resistance hybrid or 4-wire plug-in unit permits access to an 11-pin vacuum tube socket where 4-wire connections can be made to the equipment for test. This socket also provides access to *E* and *M* signaling leads. Just to the right of the plug-in resistance hybrid, an 11-pin vaccum-tube socket provides bridging point access to input and output points of the circuit as well as to cathode points in the vacuum-tube circuits. Two gain controls are shown, one for adjusting signal bias and the other for adjusting net loss. Through the use of a bridging measurement in the vacuum-tube socket shown, circuit equivalent can be adjusted without removing circuits from subscriber service by adjusting the net loss gain control potentiometer. Measurement utilizes the frequency shift signaling tone.

The positions of plug-in units in the top two sections can be seen by reference to Figs. 1 and 2. These units are carrier supplies for channel and group equipment throughout the bay, group modulating, and other common equipment in the transmission circuits, or alarm equipment. Each unit is generally of the same cubic arrangement as the channel equipment on the lower three sections but has different height, as shown in Fig. 2. The 11-pin vacuum-tube socket appears throughout for test purposes as on

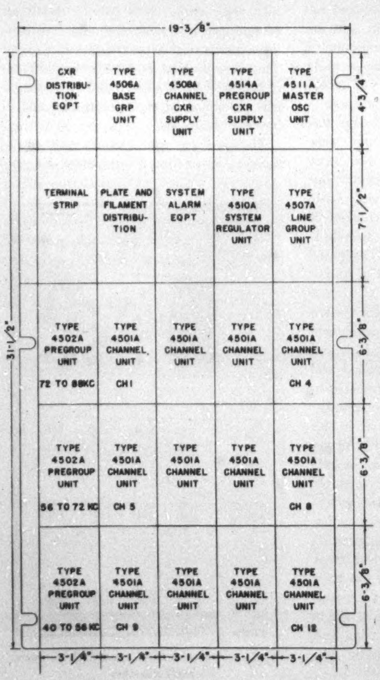

19-3/8"
CXR DISTRIBUTION EQPT
TYPE 4506A BASE GRP UNIT
TYPE 4508A CHANNEL CXR SUPPLY UNIT
TYPE 4514A PREGROUP CXR SUPPLY UNIT
TYPE 4511A MASTER OSC UNIT
4-3/4"
TERMINAL STRIP
PLATE AND FILAMENT DISTRIBUTION
SYSTEM ALARM EQPT
TYPE 4510A SYSTEM REGULATOR UNIT
TYPE 4507A LINE GROUP UNIT
7-1/2"
TYPE 4502A PREGROUP UNIT 72 TO 88KC
TYPE 4501A CHANNEL UNIT CH 1
TYPE 4501A CHANNEL UNIT
TYPE 4501A CHANNEL UNIT
TYPE 4501A CHANNEL UNIT CH 4
6-3/8"
TYPE 4502A PREGROUP UNIT 56 TO 72 KC
TYPE 4501A CHANNEL UNIT CH 5
TYPE 4501A CHANNEL UNIT
TYPE 4501A CHANNEL UNIT
TYPE 4501A CHANNEL UNIT CH 8
6-3/8"
TYPE 4502A PREGROUP UNIT 40 TO 56 KC
TYPE 4501A CHANNEL UNIT CH 9
TYPE 4501A CHANNEL UNIT
TYPE 4501A CHANNEL UNIT
TYPE 4501A CHANNEL UNIT CH 12
6-3/8"
31-1/2"
3-1/4"
3-1/4"
3-1/4"
3-1/4"
3-1/4"

Fig. 2 (left). Equipment arrangement for 12-channel type-45A terminal shelf arrangement

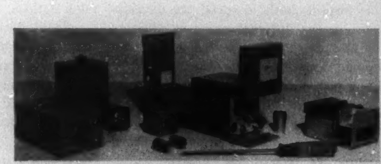


Fig. 4 (right). Channel unit with all plug-in subassemblies removed

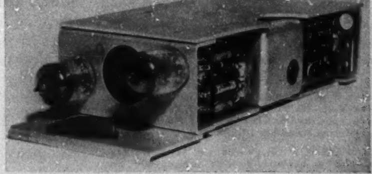


Fig. 5 (right). Compandor unit

Vo
Pt.
19

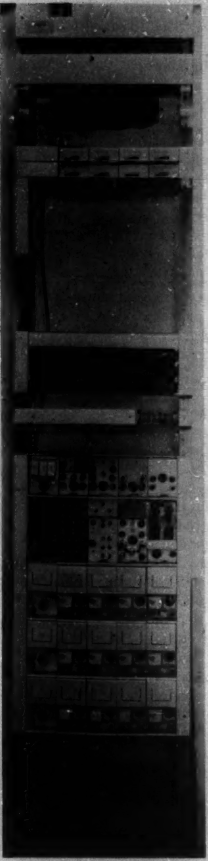


Fig. 6. Factory rack assembly of 45A system including 12-channel terminal shelf assembly and accessory equipment

the units on the lower three sections. Certain of these units are further broken down into plug-in subassemblies either for test purposes or to standardize arrangements provided at East and West terminals and for East and West direction repeaters.

The base group unit shown as the second unit from the left in the top section contains the plug-in pilot oscillators. Two versions of this are provided with one supplying pilots at 40 and 80 kc and the other at 99 and 150 kc. Change of type for East or West terminal use requires only substitution of plug-in crystals. Terminal lineup and test can be conducted on a local loop basis by removing the plug-in pilot subassembly and inserting in its jack a plug-and-cord fixture connecting to test equipment. This connects the 40- to 88-kc frequency band from the transmitting branch of the terminal back into the receiving branch.

In the carrier distribution unit at the left on the top shelf, plug connectors and jacks shown in rectangular-shaped cans afford U-link connections of the transmitting and receiver carried paths. Removal of the plugs provides test access. To the rear of the carrier distribution unit, option plugs are provided in two sockets. These enable modulator or demodulator channel carrier frequencies to be supplied at either upper or lower frequency edges of the channel bands. This plug would be used as follows:

If one type of plug in the modulator carrier supply jack were used, carriers at 8, 12, 16, and 20 kc would be provided to the channels in the three rows of channel equipment. A second type of plug inserted in this jack would provide carriers to the modulators of these channels at 12, 16, 20 and 24 kc. Optional use of two such plugs provide carriers for upper side-band or lower side-band transmission in the demodulator circuits. The third plug on the front face of the carrier distribution unit is in a jack providing access to channel and pregroup carrier supply frequencies. A cord plugged from this jack to a corresponding jack on an adjacent terminal can supply carrier to the equipment of the second terminal. This permits emergency operation or removal of carrier supply units from the inoperative terminal.

In the second section of units in Figs. 1 and 2, the alarm panel is shown. It is provided in two versions. When out-of-band signaling and dialing are used, subscriber circuits must be disconnected and subsequently made busy by alarm relays following a system failure. This feature is provided by a subassembly in the alarm panel on an optional basis. Two button keys and a light on the front of this subassembly provide a test following such system failures to determine that normal transmission has been restored to and from a distant unattended terminal, and to allow locked-up alarm relays at the distant terminal to be reset. As a further function of this alarm panel, space is provided at the center where a common oscillator unit at 3,550 cycles can be inserted in those cases where inband signaling of the type used in the nation-wide toll switching plan is required. In such cases, the signal units plugged into the individual channel circuits are replaced by considerably simplified plug-in arrangements excluding the frequency shift signal oscillator.

Fig. 7. A-c power supply unit

Fig. 8. Pole-mounted line filter

An interesting feature of the master oscillator shown on the top section at the right is the inclusion of an oven-controlled crystal housing at the rear. This oven, through thermostatic means, holds the crystal temperature to very close tolerances. Space is provided in the oven for two crystals, both of a plug-in type. A 96-kc crystal oscillator is used at all times for generation of channel and prefrequencies. The second crystal oscillator is at 187, 188, 189 or 190 kc, depending on the proper frequency allocation.

In the line group unit shown to the right in the second shelf, two versions are required, one for East-West transmission and one for West-East transmission at terminals and repeaters. The East-West and West-East directional filters are both contained in a single can on this unit. In one case half of the directional filter serves as the transmitting filter and the other half as the receiving filter. It is converted from East-West to West-East operation through insertion of a plug-in subassembly in the rear of the line group unit. This subassembly contains the equalizer for the directional filter. Two types of equalizers are needed, one for low-group and one for high-group transmission. Alternative use of two such plug-in subassemblies supplies the two optional connections needed for directional filters.

The regulator unit shown next to the line group unit is used at terminals and repeaters for both directions of transmission. A crystal pick-off filter unit for the two pilot frequencies plugs in at the top on the front face. Two types are provided. Also required are two types of slope equalizer networks for low- or high-group transmission. These plug in behind the front face. Within the plug-in line equalizer can for high-group use is a roof filter to prevent repeater sing, or radio station interference above the transmission band. The low-group line equalizer can also contains an equalization network for the system line filters.

In Fig. 3, the channel unit is shown removed from the frame with all subassemblies plugged in. Shown in Fig 4 is the channel unit with all of the subassemblies removed. The plug-in band filters and voice frequency filter are in the large rectangular cans shown to the left in Fig. 4. Two versions of band filters and two versions of voice filters are provided. One set of filters provides bandwidth in the channel from 250 to 3,100 cycles; a second set of filters provides a band extending from 200 to 3,400 cycles for those cases where the use of the wider bands is demanded. In Fig. 4 the plug-in relay appears next to the plug-in filters. Just to the right is the plug-in resistance hybrid. At the extreme right is the plug-in signal unit. The bare chassis is shown in the center with its resistance card assemblies and modulator-demodulator units.

In Fig. 5, the compandor is shown. This is an assembly $3^1/_4$ inches wide and approximately 2 inches in height. Four-wire or 2-wire operation is provided through use of the same type of plug-in options used on channel units. This is visible on the side shown in the figure. Much compactness has resulted from use of unitized can assemblies containing a number of varistors and resistance parts.

In an assembly similar to the compandor unit, ringdown converter panels are provided for conversion of 20 cycles to d-c signals. Four-wire terminating sets are provided where an external hybrid is needed with optional sizes of compromise networks.

Accessory Equipment

In Fig. 6, is shown a factory assembly containing on a wired rack the basic 12-channel terminal and other accessory panels. The 12-channel terminal shelf assembly is the same as that shown in Fig. 1. However, this completely wired rack assembly also includes jackfield arrangements and other accessory equipment. Eight ringdown converter units are shown in the two shelves above the jackfield, and an a-c power supply is shown at the top of the rack. A power distribution panel is included just above the terminal block, and a telephone set panel is mounted between the terminal shelf assembly and the jackfield.

In Fig. 7, the a-c power supply unit for the 45-terminal equipment is shown. Both plate and filament supply are obtained from the single unit. The plate supply is of 1.2-ampere capacity at 130 volts. The filament transformer supplies 7.5 amperes at 25.2 volts. Line voltage of 115 volts at 50 or 60 cycles is required. A center tap on the filament transformer allows 12.6 volts to be supplied at $7^1/_2$ amperes on either side of the secondary winding without unbalancing the transformer either magnetically or thermally. The transformer has a grain-oriented steel-ribbon type of core to save space. The plate transformer steps 115-volt input to approximately 165 volts. Taps are provided on the primary to reduce this rms voltage if a constant voltage transformer is used on the input winding and to increase this voltage in case of rectifier aging. Taps are also provided on the secondary to compensate for the rectifier and filter losses as the load current increases. By this arrangement 130-volt output is obtained of adequate accuracy for load currents from 400 to 1,200 milliamperes. The rectifier is of a bridge-connected, dry-disc selenium type using 32 3-inch square plates. The power supply filter is made up of two smoothing chokes and long-life electrolytic capacitors. The two chokes are mounted in one can to occupy minimum space. Ripple is about 1/10 per cent with 60-cycle supply frequency. Tests indicate that the external field from this power supply is extremely small because of the hypersil core construction used, and that it can be used in quite close proximity to *45A* equipment without introducing objectionable noise.

In Fig. 8, two views are shown of the line filter and its housing for pole mounting use on open-wire lines. The box is approximately 14 by 10 by $8^1/_2$ inches. It is of sand-cast construction with a hinged door for access, and mounts below the crossarm with the four studs shown. Within the box are, in addition to line filter networks, protector units of a

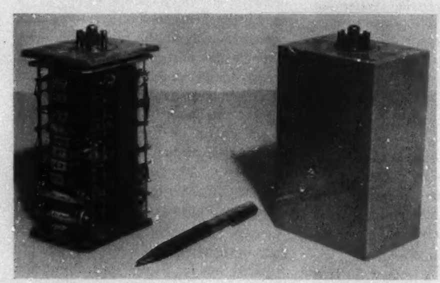
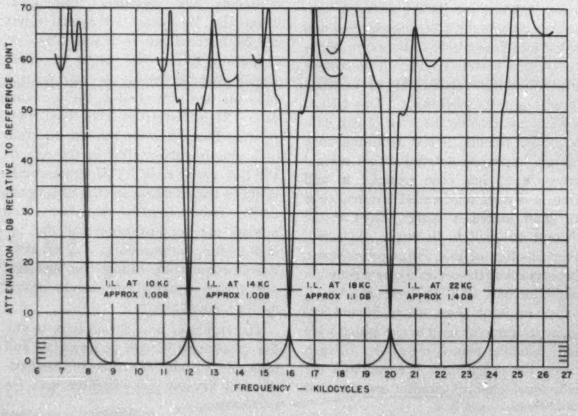
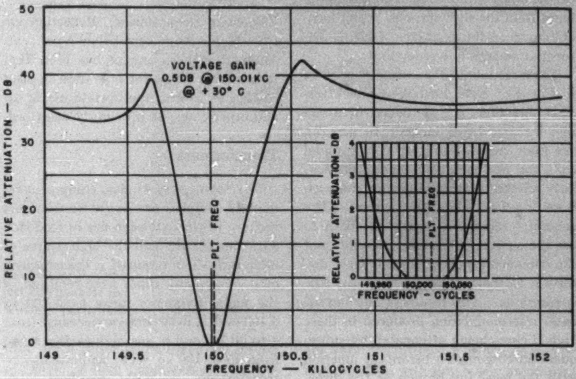


Fig. 9. Channel band filters



screw-in type, a repeat coil for connection to either 130-ohm office entrance cable or 240-ohm disc-insulated cable, and a resistance network board with strap options to make the line filter equipment match the various impedances of differently spaced open-wire line pairs. Bridle-wire connections to the line pairs and to the drop equipment are made through entrances on the bottom side of the cabinet. The hinged-door cover is gasketed to exclude driving rain and snow. The two-line filter cans are mounted within the box, with the low-pass filter can on one side and the high-pass part of the filter on the other side. In cases where separation of the carrier or voice bands is made by junction filter arrangements at the pole, two junction filters can be mounted in the two-can spaces shown within the housing. These cans are identical in size and shape to the line filter units. Junction filter arrangements are provided for separation of the *45A* system frequencies from lower carrier and voice frequencies in a 38-kc junction filter or between voice and carrier frequencies in a 3- or 5-kc junction filter.

Fig. 10. Crystal pilot pick-off filters



Filters

In Fig. 9, two views of the channel band filter are shown, both enclosed in and removed from its can. This hermetically sealed filter uses an 11-pin vacuum-tube base plug. Permalloy core coil tuned meshes are used with cores of about 3/4-inch diameter. Both modulator and demodulator band filters are contained in the same can. Mesh tuning in these filters takes advantage of the fact that no extremely great accuracy is required because all frequencies lie below 24 kc. This permits selection and grading into a small number of categories of coils wound to precision. Precision winding of polystyrene-type capacitors and a similar grading procedure into a small range of categories allows proper association of graded capacitors and coils without any precision tuning of individual meshes. A performance characteristic of this filter, is also shown in Fig. 9. As in all channels in present-day carrier equipment, a further widening and flattening of the channel characteristic is made in the voice receiving circuit of the channel. Both the high- and low-frequency part of the band are flattened.

In Fig. 10, the crystal pilot pick-off filter is shown. Two such filters are used in the system, one for the low-group pilots and one for the high-group pilots. Each plug-in filter thus contains two crystal selective filters, each filter requiring two crystals. Five-degree *X*-cut crystals are used for the three upper frequency filters. *N-T* cut crystals are used in the 40-kc filter. A transmission characteristic for the 150-kc filter is also shown in Fig. 10.

Repeaters

The repeater assembly is shown in Fig. 11. This is a shelf arrangement $7^1/_2$ inches high and of the same general

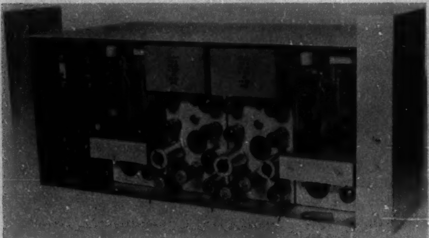


Fig. 11. Type-45A repeater shelf assembly

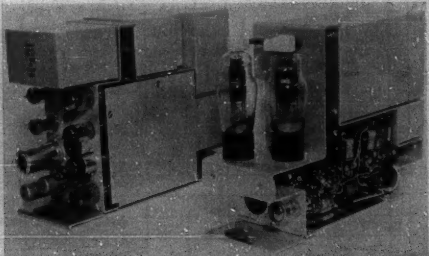


Fig. 12. Line group and regulator unit

dimensions used in the second shelf from the top in the channel terminal framework. Each direction of transmission uses a line group unit and a regulator unit which are shown in Fig. 12. The additional parts of the repeater shelf assembly are fusing, terminal block, and alarm relay arrangements. Levels of the two pilots setting the flat and slope gain of the repeater are controlled by two adjusting potentiometers on the front of the regulator unit. Level measurements can be taken without removing the equipment from the shelf through 11-pin sockets shown on the front face of the regulator unit and the line group unit.

The directional filter assembly is shown in Fig. 13. Both East-West and West-East directions are contained in a can $3^1/_4$ by $3^1/_2$ by 6 inches. In general, the filter includes both air core toroidal coils and magnetic-core coils of particularly selected permeabilities for the positions used in the filter circuit. The general principle of design has been to use the air core coils next to the line and the cores with highest Q's and most distortion farthest removed from the line connections. This results in any out-of-band intermodulation produced in these high-Q cores being attenuated by subsequent air core meshes before the junction with the receiving branch of the directional filter. Care is taken also in the design that certain of the critically tuned meshes make use of the more stable core materials.

In Fig. 14, performance curves are shown for the directional filters. The transmission frequency characteristic is shown with and without the directional filter equalizer. Great stress has been put on obtaining adequate reflection coefficients and adequate modulation performance. Impedance design has followed the methods described in the literature in recent years by Bode of the Bell Telephone Laboratories. Reflection coefficient has been kept below 5 per cent. Modulation performance has been kept to adequate levels for operation in sleet conditions by the appropriate use of air and other cores, as previously described.

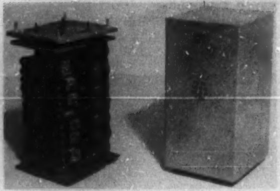


Fig. 13. Directional filters

Transformers

Development of *45*-class equipment required a complete new line of transformers because of the extensive use of feedback amplifier circuits and the high degree of miniaturization required. Transformers of a great many types were required in the audio frequency range from 200 to 3,400 cycles, in the low-carrier range from 4 to 30 kc, and in the high-carrier range from 40 to 150 kc. The audiofrequency transformers generally involved high winding impedances on tube inputs to a maximum of 600,000 ohms. Probably the major concern in output transformer design was obtaining adequate low-frequency response with the amount of plate current needed for vacuum tubes. This generally required the largest of the whole range of transformers used in the *45A* system. It has been possible to obtain the transmission performance required for all *45A* transformers with laminated cores. Multiwound, paper-insulated transformer winding was used with interleaved laminations throughout. Nickel laminations as thin as 6 mils and with dimensions as small as 5/8 by 3/4 inch were used. In a few cases in designs developed later, bobbin-type coils have been used. This has generally resulted in a considerably smaller size of transformer without impairment of audiofrequency response. Size and type of wire has been an important factor. Enamels of a newer type, able to withstand the mechanical abrasion present during winding processes were used. These do not require stripping before soldering to the permanent leads.

Throughout the design range, electrostatic shields were not necessary in most cases because of the transformers' small size. In a few cases it was necessary to use Mumetal cans for better magnetic shielding, but in most cases small extruded aluminum cans were used. These, combined with properly selected thermosetting compounds, were found to give a very satisfactory moisture seal. One of the major factors in cost reduction in the transformers resulted from a new method of attaching leads and insulating coils from the can. Two molded plastic shells completely surround the coil where it passes through the laminations. Two leads can be embedded in each end of the shells. The coil wires are wrapped on the ends of these embedded leads. Use of magnet wire insulated with semicured nylon which strips at soldering temperature allowed a solder-dip process to be used for the entire transformer, soldering all leads at once. In Fig. 15 is shown the transformer embedded in its can, the can, and the transformer alone. Tabs on the sides of the transformer can are used for mounting on the phenolic component cards. The coiled wire on the empty can is the ground connection. The transformer shown is used as an output transformer in the signaling oscillator in the channel unit.

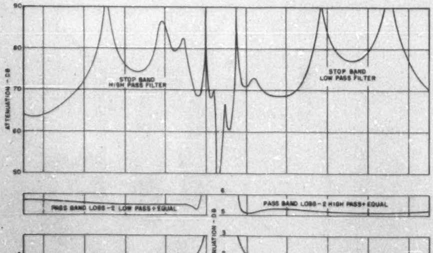


Fig. 14 (left). Directional filter performance characteristics

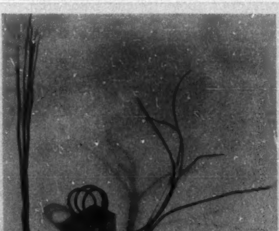


Fig. 15 (right). Typical transformer for 45-class equipment

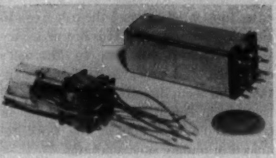


Fig. 16 (left). Modulator assembly

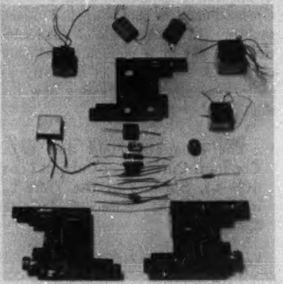


Fig. 17 (left). Phenolic component mounting cards

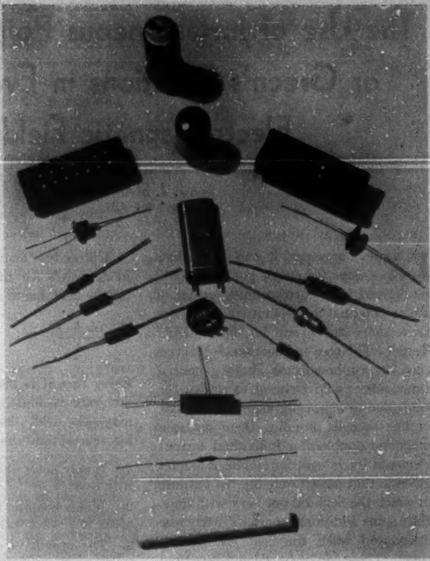


Fig. 18 (right). Typical components for 45-class equipment

Modulators

In Fig. 16, two views of a modulator from the *45A* system are shown. In one view, two transformers and the varistor assembly are shown removed from the can. In the other they are enclosed with terminal connections on a phenolic board at the top. This type of modulator is used both in the channel modulators in the 8- to 24-kc range, and in the group modulators.

In Fig. 17, the three card assemblies used in the *45A* channel unit are shown with some of the associated components removed. All of the transformers are supported on these cards by tabs at the ends, and the flexible wire leads are brought out with Radio Manufacturers Association color codes for terminal numbers. Components of the pigtail type are mounted on staples set in the cards by machine operation.

Components

Some of the *45*-class components are further shown in Fig. 18. All modulators in channel and group circuits employ germanium diode varistors arranged in the double-balanced or ring circuits. Nearly all of the filter components use wedding-ring size permalloy toroidal coils and polystyrene capacitors. Feedback amplifiers designed throughout on a 1- or 2-stage basis use miniature pentode or double-triode tubes.

Along with great stress on miniaturization, much emphasis has been put on use of good sealing compounds against humidity, on better fungus protection for tropical use, on use of longer life, higher grade performance for vacuum tubes, and on use of more stable and precise resistances. In some instances it has been found necessary to take diodes received in sealed form from manufacturers and further seal these units by additional encasement in plastic resins. Vacuum tubes are few in number in the *45A* system, and the design has not encountered any excessive heat problem. Stress on long life has influenced the design to use of special versions of the miniature *6AK5* and *2C51* type of tube. Western Electric Company versions of these tubes with 20-volt heaters (*408A* and *407A*) have been used widely in telephone plants for a number of years.

Conclusion

It is believed that the *45*-class equipment design is well adapted to present-day needs in application of carrier equipment. More and more, carrier is spreading to remote offices formerly operating only on a voice-frequency basis. In such cases, maintenance is frequently a problem, and the extensive use of plug-in assemblies easily removed and sent to more centrally located repair centers is becoming a necessity. Also, test methods where operation can be determined by bridging tests with simple test equipment are invaluable in small offices which may normally be unattended or have little skilled technical personnel. The simplest of check test can be made in these cases by ohmmeter or vacuum-tube voltmeter which take little technical training for satisfactory operation and use. A small assortment of spare components quickly plugged in, such as relays and filters, are part of this modern method of maintenance. Although development of the *45* class of terminal equipment cannot be called finished, since many improvements will come as the initially developed *45A* equipment is expanded into use in cable and on radio, it seems quite certain to the designers that the development approach from a mechanical standpoint has afforded many benefits in easing the design effort, in production, and in customer use.

Reference

1. APPLICATION AND TRANSMISSION FEATURES OF A NEW 12-CHANNEL OPEN-WIRE CARRIER SYSTEM, K. E. Appert, R. S. Caruthers, W. S. Chaskin. *AIEE Transactions*, vol. 73, pt. I, March 1954, pp. 18–27.

◆

No Discussion

The Use of Instantaneous Point Sources or Green's Functions in Evaluating Electromagnetic Fields

J. J. SMITH
MEMBER AIEE

IN a previous paper[1] it was shown how expressions for vector potential could be derived from corresponding values of scalar potential for a point source; these expressions are given in the author's Tables of Green's Functions.[2] It was then shown how these vector-potential expressions could be used to derive the fields due to uniformly magnetized disks, or due to coils, building them up from magnetic doublets or individual current elements by integration. The examples given in that paper[1] were limited to those in which the field did not vary with time, since some additional concepts need to be introduced when the field varies with time. These will be considered here, and examples given of the field due to conductors or coils in which the current varies with time.

The method followed is fairly simple, and may briefly be described as follows. To find the field due to a current in a conductor, the conductor is divided up into elements of length ds. At each of these elements, there is considered to be an instantaneous point charge Q_1 moving with a velocity **u**, with components u_x, u_y, and u_z. The scalar potential v due to this instantaneous charge can then be found for the given boundary conditions from the Tables of Green's Functions. Having obtained the scalar potential v, it is well known (see, for example, references 3 and 4) that the solution of Maxwell's equations for the vector potential **A** with components A_x, A_y, and A_z is*

$$A_x=\epsilon vu_x,\ A_y=\epsilon vu_y,\ A_z=\epsilon vu_z \tag{1}$$

The value of **A** thus obtained is integrated with respect to time, from a time t_0 when the current is started in the circuit, up to the present time. This result is then integrated over all the elements of the circuit to obtain the vector potential for the complete circuit.

A procedure analogous to this is found in Ampere's (or Biot and Savart's) law for the field due to a current element

$$H=\frac{ids\sin\theta}{r^2} \tag{2}$$

Using this, it is theoretically possible to integrate around a circuit and find the magnetic field. However, the formula has several limitations, among which are:

1. It only holds when there is no variation with time.
2. It only holds for free space and does not apply to a space with finite boundaries.
3. The formula, as written above, is in the spherical co-ordinate system. Similar expressions in other co-ordinate systems are needed.

When there is variation with time, equation 2 may be replaced by other expressions including retarded potential. An example will be given in Section 2 for a particular case involving free space, where the boundaries are at an infinite distance. When, however, the boundaries of the space are at finite distances, as in the case or cavities or wave guides, the method generally followed is to start from Maxwell's wave equation.

In the following, we start with the expression for the field due to an instantaneous point source, obtained from the Tables of Green's Functions. If the problem involves free space, the boundaries are taken to be at infinity. In this case it is shown how to derive, from the Green's Function, the field due to a current which may or may not vary with time, in a straight wire of finite or infinite length or a rectangular coil (or for any coil, if we approximate it by suitable rectangles). If a finite portion of space is involved, such as a resonant cavity or wave guide, the solution for the instantaneous point source is chosen from the Tables of Green's Functions so as to satisfy the conditions to be met at the boundaries of the cavity.

When the field due to an instantaneous point source is thus obtained, only the operation of integration is needed in the cases illustrated to meet the other conditions in the problem. It is not necessary to go back in each case to Maxwell's equations, and endeavor to find a solution from them directly which satisfies all the given conditions. This results in considerable reduction in the work required in finding such solutions.

The purpose of this paper is to show by illustrative examples how the Tables of Green's Functions may be applied to such problems. The discussion of these solutions is kept at a minimum and only covers new features when necessary so as to concentrate on showing how the solutions may be obtained.

1. The Magnetic Field due to a Constant Current I or Current I cos $\omega(t-t_0)$, Starting at the Time t_0, in a Straight Wire of Infinite Length

Let us first consider the field due a current $I\cos\omega(t-t_0)$. The first step in the solution of this problem is to find the potential at x, y, z due to an instantaneous charge Q_1 in free space at the point x_1, y_1, z_1 and at the time $t=t_1$. This will be found in the Tables of Green's Functions[2] (Section VI: Green's Function, Instantaneous Source Three Dimensions; Case III) on taking the boundaries $l_1=-\infty$, $l_2=\infty$, $m_1=-\infty$, $m_2=\infty$, $n_1=-\infty$, and $n_2=\infty$.

The solution is

$$v=g_0QX_{1p}(-\infty,\infty)Y_{1q}(-\infty,\infty)\\ Z_{1u}(-\infty,\infty)T \tag{3}$$

where

$$\begin{aligned}X_{1p}(-\infty,\infty)&=\frac{1}{\pi}\int_0^\infty\cos\alpha(x-x_1)d\alpha\\ Y_{1q}(-\infty,\infty)&=\frac{1}{\pi}\int_0^\infty\cos\beta(y-y_1)d\beta\\ Z_{1u}(-\infty,\infty)&=\frac{1}{\pi}\int_0^\infty\cos\gamma(z-z_1)d\gamma\\ T&=\sin\,[c(t-t_1)(\alpha^2+\beta^2+\gamma^2)^{1/2}]\\ &\qquad H(t,t_1)/[(\alpha^2+\beta^2+\gamma^2)/c^2]^{1/2}\end{aligned} \tag{4}$$

It seems desirable to give the solution above in the form in which it is tabulated in the Tables of Green's Functions. This solution, as explained in those tables, is written in the usual manner, putting all the integral signs in front of the right-hand side giving

---

* It was pointed out in reference 1 that there is no agreement on whether μ should be included in **A** in its definition. The notation followed here is the same as that used in reference 1 and also agrees with that used in reference 3. The notation used in reference 4 would give $\mathbf{A}=\mu\epsilon v\mathbf{u}$. However, it gives the same field intensity using $\mathbf{B}=\text{curl}\,\mathbf{A}$ instead of the equation $\mathbf{B}=\mu\,\text{curl}\,\mathbf{A}$ used here.

Paper 54-162, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing November 27, 1953.

J. J. Smith is with the General Electric Company, Schenectady, N. Y.

$$v=\frac{g_0Q}{\pi^3}\int_0^\infty\int_0^\infty\int_0^\infty \cos\alpha(x-x_1)$$
$$\cos\beta(y-y_1)\cos\gamma(z-z_1)T\,d\alpha d\beta d\gamma \quad (5)$$

However, the symbols $X_{1p}(-\infty,\infty)$, $Y_{1q}(-\infty,\infty)$, and $Z_{1u}(-\infty,\infty)$ are very convenient, and reduce the amount of writing considerably; thus they will be used in the following. In later examples, in sections 6 and 7, expressions such as $X_{1p}(v,v)$ will be obtained from the tables for the case of resonant cavities and wave guides. Summation signs rather than integrals appear in these expressions and again these summation signs are all placed in front of the resulting solution given by the tables, as in the case of the integral signs given above.

From reference 2, $H(t,t_1)$ is the unit function, so that

$$\begin{aligned} H(t,t_1)&=0 \quad \text{when } t<t_1\\ &=1 \quad \text{when } t>t_1 \end{aligned} \quad (6)$$

and thus ensures that the potential in equation 3 is zero when $t<t_1$.

Let the infinite straight wire be in the x direction so that $u_x=u$ is the velocity of the instantaneous charge and u_y, u_z are equal to zero. It follows, from equation 1

$$A_x=\epsilon vu,\ A_y=0, \text{ and } A_z=0 \quad (7)$$

Assume that a charge which varies as $Q\cos\omega(t_1-t_0)$ and is instantaneously moving with a velocity u is placed at the point x_0 at the time t_0. At a subsequent moment t_1 it will be at the point $x_1=x_0+u(t_1-t_0)$ and thus the potential due to it starting at $t=t_0$ is

$$A_x=\frac{Qu}{\pi}\int_{t_0}^{t}\int_0^\infty \cos\alpha[x-x_0-u(t_1-t_0)]$$
$$Y_{1q}(-\infty,\infty)Z_{1u}(-\infty,\infty)T\,d\alpha dt_1 \quad (8)$$

This could give the potential due to a single point charge. The potential due to a current along the path of this point charge is then found by putting a series of charges along this path, which is done by integrating from one end of the path (which may be denoted by x_3) to the other end (which may be denoted by x_4); thus the potential due to this current is given by

$$A_x=\frac{Qu}{\pi}\int_{x_3}^{x_4}dx_0\int_{t_0}^{t}\int_0^\infty$$
$$\cos\alpha[x-x_0-u(t_1-t_0)]Y_{1q}(-\infty,\infty)$$
$$Z_{1u}(-\infty,\infty)T\,d\alpha dt_1 \quad (9)$$

subject to the condition that when the length of the path is finite, $x_0+u(t_1-t_0)$ lies between x_3 and x_4, and $t_1>t_0$. This formula could be used to derive the results in the following.

The problem, however, can be looked at from another point of view as follows. Consider any given point x_1: at a given time, an instantaneous point charge is placed there. Although this point charge moves on as indicated above, another point charge moves to this point, and since there is no way of identifying point charges, this point charge may be considered to take the place of the previous one. Thus the process may be regarded as being one of placing an instantaneous point charge at a given point x_1, with a given velocity u at the initial time t_0 and maintaining it there. Then, by integration from x_3 to x_4 as before, the field due to the current in the circuit is found to be

$$A_x=\frac{Qu}{\pi}\int_{x_3}^{x_4}dx_0\int_{t_0}^{t}\int_0^\infty \cos\alpha(x-x_0)$$
$$Y_{1q}(-\infty,\infty)Z_{1u}(-\infty,\infty)T\,d\alpha dt_1 \quad (10)$$

This latter formula is much simpler to use than equation 9, and from physical considerations must give the same result as equation 9. To prove this mathematically is quite involved, since it is necessary to take account of the conditions stated following equation 9, which correspond to the departure and arrival of charges at each end of the circuit when this equation is used. When, however, the field due to a single charge is required, this equation must be used. Since the problems to follow deal with the determination of the field due to a current in a specified length of circuit, equation 10 will be used. It is convenient to do the integration with respect to t_1 first, and thus using only this part in equation 10, the potential due to an instantaneous charge Q moving with velocity u placed at x_0 at the time t_0 and maintained there is

$$A_{1x}=QuX_{1p}(-\infty,\infty)Y_{1q}(-\infty,\infty)$$
$$Z_{1u}(-\infty,\infty)\int_{t_0}^{t}T\cos\omega(t-t_1)dt_1$$
$$=QuX_{1p}(-\infty,\infty)Y_{1q}(-\infty,\infty)$$
$$Z_{1u}(-\infty,\infty)T_1 \quad (11)$$

where

$$T_1=\frac{c^2}{c^2\delta^2-\omega^2}[\cos\omega(t-t_0)-$$
$$\cos c\delta(t-t_0)]H(t,t_0) \quad (12)$$

$$\delta^2=\alpha^2+\beta^2+\gamma^2 \quad (13)$$

From this the vector potential due to the current in a wire of infinite length is obtained by integrating with respect to x_1 from $-\infty$ to ∞, obtaining

$$A_{2x}=\int_{-\infty}^{\infty}dx_1A_{1x}=\frac{Qu}{\pi}\int_{-\infty}^{\infty}dx_1\int_0^\infty d\alpha$$
$$\cos\alpha(x-x_1)Y_{1q}(-\infty,\infty)$$
$$Z_{1u}(-\infty,\infty)T_1 \quad (14)$$

where T_1 is a function of α. Using Fourier's theorem, this becomes

$$A_{2x}=QuY_{1q}(-\infty,\infty)Z_{1u}(-\infty,\infty)$$
$$(T_1)_{\alpha=0} \quad (15)$$

Write

$$(T_1)_{\alpha=0}=T_2 \quad (16)$$

$$\nu^2=\beta^2+\gamma^2 \quad (17)$$

Then

$$T_2=\frac{c^2}{c^2\nu^2-\omega^2}[\cos\omega(t-t_0)-$$
$$\cos c\nu(t-t_0)]H(t,t_0) \quad (18)$$

so that the vector potential A_{2x} due to the current $\cos\omega(t-t_0)$ in the straight wire of infinite length is given by

$$A_{2x}=QuY_{1q}(-\infty,\infty)Z_{1u}(-\infty,\infty)T_2 \quad (19)$$

where $Y_{1q}(-\infty,\infty)$ and $Z_{1u}(-\infty,\infty)$ are given by equation 4 and T_2 is given by equation 18.

Here it may be noted that A_{2x} may be divided into two terms A_{2x1} and A_{2x2} by writing

$$A_{2x}=A_{2x1}-A_{2x2} \quad (20)$$

where

$$A_{2x1}=QuY_{1q}(-\infty,\infty)Z_{1u}(-\infty,\infty)$$
$$\cos\omega(t-t_0)H(t,t_0)c^2/(c^2\nu^2-\omega^2) \quad (21)$$

$$A_{2x2}=QuY_{1q}(-\infty,\infty)Z_{1u}(-\infty,\infty)$$
$$\cos c\nu(t-t_0)H(t,t_0)c^2/(c^2\nu^2-\omega^2) \quad (22)$$

Here the term A_{2x1} involves the term $\cos\omega(t-t_0)$, so that it corresponds to the impressed frequency ω. This might then be regarded as the "steady-state" term. A_{2x2} involves the term $\cos c\nu(t-t_0)$. Since $\nu^2=\beta^2+\gamma^2$, and each of these are variables of integration extending from 0 to ∞ this term is not periodic, but changes with time, and thus corresponds to what might be called the "transient" term. It ensures that the vector potential is zero at a distance $r=(y^2+z^2)^{1/2}$ from the wire when $r>c(t-t_0)$, so that the electromagnetic wave has not arrived at the point.

When the current in the wire is constant, the magnetic field may be obtained from equations 18 and 19 by putting $\omega=0$, giving

$$A_{3x}=QuY_{1q}(-\infty,\infty)Z_{1u}(-\infty,\infty)T_3 \quad (23)$$

where

$$T_3=\frac{1}{\nu^2}[1-\cos c\nu(t-t_0)]H(t,t_0) \quad (24)$$

and $Y_{1q}(-\infty,\infty)$ $Z_1u(-\infty,\infty)$ are given in equation 4.

Here again, note that A_{3x} might be divided into two parts, A_{3x1} and A_{3x2}, where A_{3x1} corresponds to the "steady state," where there is no variation with time, and A_{3x2} corresponds with the "transient" term. It may also be noted that in going from the constant-current case to that of a sinusoidal current, we not only add the term cos $\omega(t-t_0)$, but also modify the factor in front from $1/\nu^2$ to $c^2/(c^2\nu^2-\omega^2)$. This will be discussed further later.

This can be shown to give the well-known result for the field due to a constant current in a straight wire of infinite length, by dropping the "transient term" and eliminating the term $H(t,t_0)$, since the current is assumed to exist for an infinite time, obtaining

$$A_{3z1}=QuY_{1q}(-\infty,\infty)Z_{1u}(-\infty,\infty)/\nu^2 \quad (25)$$

By using a result given by Van der Pol and Bremmer[5] that

$$1/r=X_{1p}(-\infty,\infty)Y_{1q}(-\infty,\infty) Z_{1u}(-\infty,\infty)/\delta^2 \quad (26)$$

it can be shown that

$$A_{3z1}=\log\,[(y-y_1)^2+(z-z_1)^2]+C \quad (27)$$

when C is a constant of integration independent of y and z. This is, of course, the well-known result.

It may also be well to point out that in the solution corresponding to a current I cos $\omega(t-t_0)$ it has been assumed that the current is constant along the line. This is a good approximation when the frequency is low and the line is not too long. But as the frequency or the length of the line is increased, the voltage along the line causes charging or displacement currents, with the result that the current is no longer constant along the line. Thus the above solution has to be modified to allow for the variation in current along the line. This leads to transmission line theory and will be considered in Section 3 when the field due to a finite line is discussed.

2. Derivation of Ampere's Law and an Extension to Variation with Time

Ampere's law, referred to in equation 2, can be readily derived from the Tables of Green's Functions. Equations 3 and 4 give the field due to an instantaneous point charge. The field due to a point charge placed at the point at the time t_0 and maintained there is obtained by integrating with respect to t from t_0 to t, obtaining

$$\begin{aligned} v&=g_0QX_{1p}(-\infty,\infty)Y_{1q}(-\infty,\infty) Z_{1u}(-\infty,\infty)\int_{t_0}^{t}T\,dt_1\\ &=QX_{1p}(-\infty,\infty)Y_{1q}(-\infty,\infty) Z_{1u}(-\infty,\infty)[1-\cos c\delta(t-t_0)] H(t,t_0)/\delta^2 \end{aligned} \quad (28)$$

It will be noted that, as before, the term in square brackets consists of a term 1, which is independent of time and a term cos $c\delta(t-t_0)$, which depends on the time. The latter term may be considered as the "transient" and if it died away with time (which it doesn't unless dissipation is introduced in the equation), then the "steady-state" condition would be given by the other term. This other term is the one actually obtained if the problem is solved from the Tables of Green's Functions in the case when there is no variation with time. Thus, in this case, when there is no variation with time

$$v=QX_{1p}(-\infty,\infty)Y_{1q}(-\infty,\infty) Z_{1u}(-\infty,\infty)/\delta^2 \quad (29)$$

It is known[5] that the right-hand side of this equation is equal to Q/r. This result is also evident, since it is the scalar potential due to a point charge in free space, which is Q/r. Now let the point charge be moving with an instantaneous velocity u whose components along the axes are $u\ dx/ds$, $u\ dy/ds$ and $u\ dz/ds$. Then, from equation 1 the vector potential is given by

$$A_x=\frac{Qu}{r}\frac{dx}{ds};\quad A_y=\frac{Qu}{r}\frac{dy}{ds};\quad A_z=\frac{Qu}{r}\frac{dz}{ds} \quad (30)$$

which can be checked with the formula for vector potential given by Jeans[6] on putting $Qu=i$ and noting that the foregoing refers to an element of the circuit. The vector potential for the complete circuit is obtained by integrating around it with respect to s, which integral is included in Jeans' result. In paragraph 497 which follows the foregoing, Jeans shows how Ampere's law can be derived from this, and it is not necessary to repeat this here. This derivation of Ampere's law should help make clear the limitations which apply to it and were referred to in the introduction. It gives the field due to an element of current ds which does not vary with time and which is in free space, so that the boundaries of the space are at an infinite distance.

When there is variation with time, it was mentioned in section 1 that other methods, including retarded potential, may be used. As an example of one of these methods, consider an instantaneous charge which varies as Q cos $\omega(t-t_0)$ which is placed at the point x_0 at the time t_0 and is maintained there. The vector potential due to it has already been given in equation 11. If in this equation, the term cos $c\delta(t-t_0)$ is considered a "transient term," which in a practical case would die down, and can thus be omitted as in the steady-state case above, the result is

$$A_{1x}=c^2QuX_{1p}(-\infty,\infty)Y_{1q}(-\infty,\infty) Z_{1u}(-\infty,\infty)\cos\omega(t-t_0) H(t,t_0)/(c^2\delta^2-\omega^2) \quad (31)$$

But from reference 5, this is equal to

$$\begin{aligned} A_{1x}&=Qu\cos(\omega r/c)\cos\omega(t-t_0)H(t,t_0)/r\\ &=Qu[\cos\omega(t-t_0+r/c)+ \cos\omega(t-t_0-r/c)]H(t,t_0)/r \end{aligned} \quad (32)$$

If we are only interested in waves going out from the source, the first term being a converging wave does not apply. Thus taking only the second term

$$A_{1x}=Qu\cos\omega(t-t_0-r/c)H(t,t_0)/r \quad (33)$$

and this, with appropriate changes in notation, is the result given by Schelkunoff[7] when there is spherical symmetry and variation with time. It may be interesting to note that the term $c^2\delta^2-\omega^2$ in the denominator of equation 31 does not give rise to any infinite values in this case. This is, of course, due to the operations of integration which are contained in the symbols $X_{1p}(-\infty,\infty)$, etc., given in equation 4. In the case of finite boundaries, however, when terms such as X_{1p}, (v,v), involving summation, enter, this corresponding factor $c^2\delta^2-\omega^2$ does give rise to infinite values, as will be discussed in sections 7 and 8 in considering resonant cavities and wave guides.

3. The Magnetic Field due to a Current I cos ωt in a Straight Wire of Finite Length

In this case, proceed in the same manner as in the previous example and obtain equation 11 which gives the vector potential due to an instantaneous charge Q cos $\omega(t-t_0)$ which is moving with an instantaneous velocity u and is placed at the point x_1, y_1, z_1 at the time $t=t_0$ and maintained there. For the finite length of wire extending from x_2 to x_3, the vector potential is then obtained as

$$A_{4x}=\int_{x_2}^{x_3}A_{1x}\,dx_1=Qu\int_{x_2}^{x_3}dx_1X_{1p}(-\infty,\infty) Y_{1q}(-\infty,\infty)Z_{1u}(-\infty,\infty)T_1 \quad (34)$$

which is the vector potential due to a current I cos ωt starting at the time $t=t_0$ in a straight wire in free space extending from $x=x_3$ to $x=x_2$.

An antenna is essentially a wire in space carrying current. However, the current in an antenna not only varies with time, but also its maximum value at any point may vary with the distance along the antenna. This is due to the fact that, in the case of an antenna, a voltage is applied to it at a given point. The current at any point is thus equal to displacement current propagated through space from that point, due to its being at a given voltage, plus the current which passes that point to feed the rest of the antenna. The determination of the distribution of currents along the antenna may be approximated by methods sinilar to transmission-line theory. By proceeding in this manner, the determination of the field due to an antenna may be divided into two parts:

1. The magnitude of the current along the antenna;
2. The determination of the field due to the current distribution.

If it can be determined that the current distribution in the antenna is given by

$$I=I_1 f(x_1) \tag{35}$$

Then the vector potential due to this distribution of current is from equation 11

$$A_{5x}=\int_{x_2}^{x_3} A_{1x} f(x_1)dx_1 = Qu\int_{x_2}^{x_3} X_{1p}(-\infty,\infty) \\ Y_{1q}(-\infty,\infty)Z_{1u}(-\infty,\infty)f(x_1)T_1 dx_1 \tag{36}$$

By treating this equation in the same manner as equation 31, making the same assumptions that the "transient term" in T_1 can be omitted, and also that only diverging waves will be considered, it is found that it can be written in the form[5]

$$A_{5x}=Qu\int_{x_2}^{x_3} \cos\omega(t-t_0-r/c)f(x_1) \\ H(t,t_0)dx_1/r \tag{37}$$

and with suitable changes in notation this is the same result as that given by Schelkunoff[8] and Stratton.[9]

It may be well to show how easily a slightly different problem may be solved. Instead of taking the above condition that the antenna is in free space, let it be in the x direction at a distance z_1 above the earth which is the plane $z=0$, and let the earth be taken as at zero potential. Then it is found from the tables that $Z_{1u}(-\infty,\infty)$ is to be replaced by $Z_{1u}(v,\infty)$ to give the solution corresponding to the terminal condition $v=0$ when $z=0$. Thus the vector potential for this case becomes

$$A_{6x}=\int_{x_2}^{x_3} X_{1p}(-\infty,\infty)Y_{1q}(-\infty,\infty)\times \\ Z_{1u}(v,\infty)f(x_1)T_1 dx_1 \tag{38}$$

where

$$Z_{1u}(v,\infty)=\frac{2}{\pi}\int_0^\infty d\gamma \sin\gamma z \sin\gamma z_1 \tag{39}$$

and $X_{1p}(-\infty,\infty)$, $Y_{1q}(-\infty,\infty)$ and T_1 are as given previously in equations 4 and 12

This division of the antenna problem into two parts:

1. Determination of the current distribution $f(x)$
2. Determination of the field from the current distribution

is followed by many authors. Frequently a sinusoidal current distribution is assumed, and this works well in many cases. If it is desired to improve on this, an integro-differential equation can be found for $f(x)$, which is given by Schelkunoff[10], and having thus determined $f(x)$, the solution proceeds in the usual manner. This will not be considered further here except to suggest that in this integro-differential equation, the function $G(\zeta-z,a)$, which is a Green's function, may often be put into a form in which it is easier to solve the equation by using the Tables of Green's Functions to express it as a Fourier series.

4. The Magnetic Field due to a Current I cos ωt in a Rectangular Coil

Consider a rectangular coil in free space in the plane $x=x_1$, whose sides are at $y=a$, $y=-a$, $z=b$, $z=-b$. Let there be a current I cos ωt in it; the resulting magnetic field may be obtained by the following procedure.

We proceed exactly as in the previous examples, and find the vector potential A_{1x} due to a charge which varies as cos $\omega(t-t_0)$, is moving with an instantaneous velocity u, and is placed at the point x_1, y_1, z_1 at the time t_0 and maintained there. The result is given in equation 11.

In the previous example, this was integrated with respect to x_1 from $-\infty$ to ∞ to give the field due to an infinite straight wire. In the case of the coil with finite sides, we proceed as follows:

For the coil side at $y=a$, put $y_1=a$ in equation 11. Then integrate with respect to z_1 from $z_1=-b$ to $z_1=b$. This gives the contribution to the magnetic field due to the side at $y=a$. For the coil side at $y=-a$, put $y_1=-a$ in equation 11 and again integrate with respect to z_1 from $z_1=-b$ to $z_1=b$. Subtract this integral from the previous one, since the current in it is in the opposite direction. Perform a similar operation for the coil sides at $z-b=$ and $z=b$, and add to the previous result using the proper signs and the field due to the coil is obtained.

The details will not be worked out here, since they are very similar to the results for a rectangular coil when there is no variation with time, as worked out in a previous paper.[11]

5. The Relation of Vector Potential to the Terminology Used in Resonant Cavity and Wave Guide Theory

Since the terminology of resonant cavities and wave guides refer to the electric and magnetic fields and the tables we wish to use are given in terms of potential, it is first necessary to translate one type of language into the other to understand what types of potential and vector potential give rise to **E** and **H** waves.

Let it be assumed that the direction to which the properties of the waves are referred is the z direction. Thus, if E_z is zero, the wave is a transverse electric (TE) or **H** wave. If H_z is zero, the wave is a transverse magnetic (TM) or **E** wave.

Consider the general case where the charge is moving with a velocity u which has the components u_x, u_y, and u_z. Then Maxwell's equations are

$$\begin{aligned} &\nabla^2 v-\partial^2 v/c^2\partial t^2=-\rho/\epsilon \\ &\nabla^2 A_x-\partial^2 A_x/c^2\partial t^2=-\rho u_x \\ &\nabla^2 A_y-\partial^2 A_y/c^2\partial t^2=-\rho u_y \\ &\nabla^2 A_z-\partial^2 A_z/c^2\partial t^2=-\rho u_z \end{aligned} \tag{40}$$

From these, v and A can be found from the tables when the boundary conditions are given. Now consider the case where u_x and u_z are zero, so that A_x and A_z are zero. The equations for the electric and magnetic field then become

$$\begin{aligned} &E_x=-\partial v/\partial x && H_x=-\partial A_y/\partial z \\ &E_y=-\partial v/\partial y-\mu\partial A_y/\partial t && H_y=0 \\ &E_z=-\partial v/\partial z && H_z=\partial A_y/\partial_x \end{aligned} \tag{41}$$

On the other hand, if u_x and u_y are equal to zero, A_x and A_y are equal to zero, and the equations for the electric and magnetic field become

$$\begin{aligned} &E_x=-\partial v/\partial x && H_x=\partial A_z/\partial y \\ &E_y=-\partial v/\partial y && H_y=-\partial A_z/\partial x \\ &E_z=-\partial v/\partial z-\mu\partial A_z/\partial t && H_z=0 \end{aligned} \tag{42}$$

It will be noted that in the second case, where the charges move only in the z direction, $H_z=0$, as would be expected, and this corresponds with the definition of a TM or **E** wave, since there is a component of **E** in the z direction. In the first case there is a component of **H** in the z direction, but there is also a component of **E** in the z direction, namely, $E_z=-\partial v/\partial z$. However, it is seen from the equations for v and A_y that v is proportional to ρ/ϵ and A_y is proportional to ρu_y. It is possible, therefore, to have u_y very large, so that the product ρu_y is large compared with ρ/ϵ. We may, therefore, let $\rho u_y=i_y'$, the current density in the y direction, which can be taken to remain finite, although it may be assumed that $\rho\to 0$ and thus $v\to 0$. Under the latter assumption, which seems to correspond with assumptions as to **E** and **H** in resonant-cavity and wave-guide theory, the first equations become those corresponding to a current in the y direction, giving an **H** or TE wave in the form

$$\begin{aligned} &E_x=0 && H_x=-\partial A_y/\partial z \\ &E_y=-\mu\partial A_y/\partial t && H_y=0 \\ &E_z=0 && H_z=\partial A_y/\partial x \end{aligned} \tag{43}$$

The second set, with charges moving in the direction of the z axis, corresponds to an **E** wave or TM wave in the form

$$\begin{array}{ll} E_x=0 & H_x=\partial A_z/\partial y \\ E_y=0 & H_y=-\partial A_z/\partial x \\ E_z=-\mu\partial A_z/\partial t & H_z=0 \end{array} \qquad (44)$$

and obviously by comparison with these, if the charges only move in the x direction

$$\begin{array}{ll} E_x=-\mu\partial A_x/\partial t & H_x=0 \\ E_y=0 & H_y=\partial A_x/\partial z \\ E_z=0 & H_z=-\partial A_x/\partial y \end{array} \qquad (45)$$

When the given direction is the z direction, the first (equations 43) and third (equations 45) of these groups are generally combined by adding together, since it will be noted that in each of them $E_z=0$, and thus the resultant of adding them corresponds to a TE wave. We shall not combine them, however, as it is simpler in practice to work them out separately, and then add the final result.

Therefore, in the following we shall omit the electric intensity derived from the scalar potential given by the terms $\partial v/\partial x$, $\partial v/\partial y$, and $\partial v/\partial z$, on the basis that these are negligible in usual cases compared with $\mu\partial A_x/\partial t$, etc. However, these electric intensities can be easily obtained, and should not be entirely overlooked in future developments where they may prove helpful.

Thus the solution of the field in any given cavity is reduced to the determination of the three vector potentials A_x, A_y, and A_z. In order to obtain these vector potentials from the Tables of Green's Functions, it is necessary to know what the boundary conditions are. We shall take the vector potential corresponding to suitable boundary conditions in the following, and show that these conditions agree with those usually taken in terms of **E** and **H** in the usual approach.

6. Resonant Cavities

A rectangular box-like cavity is excited by a current $I\cos\omega t$ passing through it in the z direction. The resulting electric and magnetic fields inside the cavity may be obtained by the following procedure.

Let the sides of the cavity be given by $x=0$, $x=a$, $y=0$, $y=b$, $z=0$ and $z=s$. Then, from the tables, the instantaneous Green's function can be found corresponding to the solution of the wave equation for a nondissipative medium, for an instantaneous charge placed at the point x_1, y_1, z_1 at the time t_1, corresponding to the equations

$$\nabla^2v-\partial^2v/c^2\partial t^2=-\rho/\epsilon$$

and

$$\nabla^2A_z-\partial^2A_z/c^2\partial t^2=-\rho u_z \qquad (46)$$

when suitable boundary conditions are known. The current in the z direction produces a TM wave. The boundary conditions for a TM wave will be found in any textbook to be those given in equation 54 in terms of **E** and **H**. We need the corresponding boundary conditions in terms of A, which can be shown to be $A_z=0$ at the x and y boundaries and $\partial A_z/\partial z$ at the z boundaries. From the tables, it is found that for an instantaneous charge moving in the z direction

$$A_z=vu_z=Qu_zcX_{1p}(v,v)Y_{1q}(v,v)Z_{1s}(v',v')\times \\ \sin c\delta(t-t_1)H(t,t_1)/\delta \qquad (47)$$

$$X_{1p}(v,v)=\frac{2}{a}\sum_p \sin\alpha_px\sin\alpha_px_1$$

$$Y_{1q}(v,v)=\frac{2}{b}\sum_q \sin\beta_qy\sin\beta_qy_1$$

$$Z_{1s}(v',v')=\frac{2}{d}\sum_s \cos\gamma_sz\cos\gamma_sz_1 \qquad (48)$$

where $\alpha_p=p\pi/a$, $\beta_q=q\pi/b$, $\gamma_s=s\pi/d$; p, q and s take the values 0, 1, 2, 3 . . . and $\delta^2=\alpha_p{}^2+\beta_q{}^2+\gamma_s{}^2$. Here again, as explained in section 1, in writing the result in the usual manner, the summation signs in equation 48 are placed in front of the right-hand side of equation 47.

Since the current is given by $I\cos\omega(t-t_0)$, this corresponds to charges varying as $Q\cos\omega(t-t_0)$, and an integration with respect to y corresponding to the length of circuit. Assume the charges have been flowing from the time t_0 to t; then, the vector potential is obtained by integrating $A_z\cos\omega(t_1-t_0)$ with respect to t_1 from t_0 to t, obtaining

$$A_{z1}=Qu_zc^2X_{1p}(v,v)Y_{1q}(v,v)Z_{1s}(v',v')\times \\ \left[\frac{\cos\omega(t-t_0)-\cos c\delta(t-t_0)}{\omega^2-c^2\delta^2}\right]H(t,t_1) \qquad (49)$$

The last term inside the square bracket corresponds to the initial "transient," and in any cavity or wave guide, with even a small amount of dissipation, will die away with time so that the "steady state" A_{z2} is given by

$$A_{z2}=Qu_zc^2X_{1p}(v,v)Y_{1q}(v,v)Z_{1s}(v',v')\times \\ \cos\omega(t-t_0)/(\omega^2-c^2\delta^2) \qquad (50)$$

which is the "steady-state" vector potential in the cavity, due to a current element $I\cos\omega(t-t_0)$ in the z direction at the point x_1, y_1, z_1.

This result is in the form of a triple infinite series of the Fourier type. However, it will also be noted that A_z contains in the denominator the quantity $\omega^2-c^2\delta^2$. By properly choosing the dimensions of the box, this quantity may be made as small as we please, and if it were not for dissipation in the cavity, A_{z2} would become infinite. This case, however, is the one that is of interest, and to study it we need only the one term of the series for which $\omega^2-c^2\delta^2=0$, since all the other terms will be small compared to it. Assume that the size of the cavity is such that the relation holds for given values of p, q, and s. The summation signs in $X_{1p}(v,v)$, $Y_{1q}(v,v)$, and $Z_{1u}(v',v')$ may be dropped. Using p, q, and s for these specific values, A_{z2} becomes

$$A_{z2}=\frac{8Quc^2}{abd(\omega^2-c^2\delta^2)}\sin\frac{p\pi x}{a}\sin\frac{p\pi x_1}{a}\times \\ \sin\frac{q\pi y}{b}\sin\frac{q\pi y_1}{b}\cos\frac{s\pi z}{d}\times \\ \cos\frac{s\pi z_1}{d}\cos\omega(t-t_0) \qquad (51)$$

and from this, if we write

$$\frac{8Quc^2}{abd(\omega^2-c^2\delta^2)}\sin\frac{p\pi x_1}{a}\sin\frac{p\pi y_1}{b}\times \\ \cos\frac{s\pi z}{d}=K \qquad (52)$$

Then, from equation 44

$$E_z=K\omega\sin\frac{p\pi x}{a}\sin\frac{q\pi y}{b}\cos\frac{s\pi z}{d}\sin\omega(t-t_0)$$

$$H_x=\frac{q\pi K}{b}\sin\frac{p\pi x}{a}\cos\frac{q\pi y}{b}\cos\frac{s\pi z}{d}\times \\ \cos\omega(t-t_0) \qquad (53)$$

$$H_y=-\frac{p\pi K}{a}\cos\frac{p\pi x}{a}\sin\frac{q\pi y}{b}\cos\frac{s\pi z}{d}\times \\ \cos\omega(t-t_0)$$

and these are the usual forms for transverse magnetic TM waves or **E** waves. We can now see that the values for the boundary conditions with respect to A_z chosen above correspond with the following set of equations

$$\begin{array}{ll} E_z=0 & \text{when } x=0, \\ \partial H_x/\partial y=0 & \text{when } y=0, \\ \partial H_y/\partial x=0 & \text{when } x=0, \\ E_z=0 & \text{when } x=a, \\ \partial H_x/\partial y=0 & \text{when } y=b, \\ \partial H_y/\partial x=0 & \text{when } x=a, \end{array} \qquad (54)$$

and these are the relations taken for E_z, H_x, and H_y in the usual approach.

The vector potentials A_x and A_y due to a current $I\cos\omega(t-t_0)$ in the x and y directions, respectively, can be written down by comparison

$$A_x=K_1\cos\frac{p\pi x}{a}\sin\frac{q\pi y}{b}\sin\frac{s\pi z}{d}\cos\omega(t-t_0)$$

$$A_y=K_2\sin\frac{p\pi x}{a}\cos\frac{q\pi y}{b}\sin\frac{s\pi z}{d}\cos(\omega t-t_0) \qquad (55)$$

From these the values of **E** and **H** can be obtained, using equations 43 and 45. These two latter results are generally added together to give the complete expression for the TE field. The results may then be compared with those given by Schelkunoff.[12]

7. Wave Guides

The wave-guide theory is quite similar to the theory of the resonant cavity just discussed. It uses another form of the solution of the wave equation, which can be readily derived from the solution for a resonant cavity above. Write equation 51 in the form

$$A_z=\frac{2Qu_z}{d}\sum_{s} X_{1p}(v,v)\,Y_{1q}(v,v)\times$$
$$\cos w(t-t_0)\,\frac{\cos\gamma_s z\cos\gamma_s z_1}{\omega^2-c^2\delta^2}\qquad(56)$$

where $X_{1p}(v,v)$ and $Y_{1q}(v,v)$ are given by equation 48 and γ_s and δ are given in the lines following equation 48. This equation for A_z can be summed with respect to s giving

$$\sum_{s}\frac{2\cos\gamma_s z\cos\gamma_s z_1}{d[\omega^2-c^2(\alpha_p^2+\beta_q^2)-c^2\gamma_s^2]}$$
$$=\frac{\cosh\gamma z\cosh\gamma(d-z_1)}{\gamma\sinh\gamma d}\times$$
$$[1-H(z,z_1)]+\frac{\cosh\gamma z_1\cosh\gamma(d-z)}{\gamma\sinh\gamma d}\times$$
$$H(z,z_1)\qquad(57)$$

where $c^2\gamma^2=\omega^2-c^2(\alpha_p^2+\beta_q^2)$. This result may be readily obtained by expanding the right-hand side of this equation in partial fractions.

Since the source exciting the wave guide is located at z_1 to study the propagation along the wave guide, it is usual to take only the coefficient of $H(z,z_1)$ in the foregoing, since it gives the vector potential when $z>z_1$. Thus the vector potential becomes, when $z>z_1$

$$A_z=Qu_zX_{1p}(v,v)\,Y_{1q}(v,v)\cos\omega(t-t_0)\times$$
$$\frac{\cosh\gamma z_1\cosh\gamma(d-z)}{\gamma\sinh\gamma d}\qquad(58)$$

We note, of course, that if $\gamma=0$, the vector potential tends to ∞, as in the case of the resonant cavity. However, if $\gamma\neq0$ there are two cases to be considered

1. If $\omega^2-c^2(\alpha_p^2+b_q^2)>0$,
then γ is real (59)
2. If $\omega^2-c^2(\alpha_p^2+\beta_q^2)<0$,
then γ is imaginary (60)

If γ is real, the type of propagation is most readily seen by making the length of the line d infinite, giving

$$A_z=Qu_zX_{1p}(v,v)\,Y_{1q}(v,v)\cos\omega(t-t_0)\times$$
$$\frac{e^{-\gamma z}\cosh\gamma z_1}{\gamma}\qquad(61)$$

Here the term $e^{-\gamma z}$ falls off exponentially with z, and rapidly reduces the fields to a low value, except when γ is small. Thus only a negligible amount of energy would be transmitted over such a wave guide The region corresponding to these values of γ is called the region of high attenuation.

When γ is imaginary, however, the hyperbolic terms in z are replaced by sinusoidal functions, and writing $\gamma_a=i\gamma$, the vector potential becomes (writing the values of $X_{1p}(v,v)$ and $Y_{1q}(v,v)$ in full)

$$A_z=\frac{4Qu_z}{ab}\sum_{p}\sum_{q}\sin\alpha_p x\sin\alpha_p x_1\times$$
$$\sin\beta_q y\sin\beta_q y_1\cos\omega(t-t_0)\times$$
$$\frac{\cos\gamma_a z_1\cos\gamma_a(d-z)}{\gamma_a\sin\gamma_a d}\qquad(62)$$

For comparison with the expressions given in the literature, write

$$K=\frac{4Qu_z}{ab}\sum_{p}\sum_{q}\sin\alpha_p x_1\sin\beta_q y_1\times$$
$$\frac{\cos\gamma_a z_1}{\gamma_a\sin\gamma_a d}\qquad(63)$$

Then

$$A_z=K\sin\frac{p\pi x}{a}\sin\frac{q\pi y}{b}\cos\gamma_a(d-z)\times$$
$$\cos\omega(t-t_0)$$

$$E_z=K\omega\sin\frac{p\pi x}{a}\sin\frac{q\pi y}{b}\cos\gamma_a(d-z)\times$$
$$\sin\omega(t-t_0)$$

$$H_x=\frac{Kq\pi}{b}\sin\frac{p\pi x}{a}\cos\frac{q\pi y}{b}\cos\gamma_a(d-z)\times$$
$$\cos\omega(t-t_0)$$

$$H_y=-\frac{Kp\pi}{a}\cos\frac{p\pi x}{a}\sin\frac{q\pi y}{b}\cos\gamma_a(d-z)\times$$
$$\cos\omega(t-t_0)\qquad(64)$$

which may be compared with those given by Southworth,[13] who does not write explicitly the terms involving time.

It will be noted that in the case of the resonant cavity and wave guide, the fields derived were those due to a point source. To obtain the field due to the current in a finite length of wire, integrate over the length of the wire, as in the previous examples.

8. Other Co-ordinate Systems

The Tables of Green's Functions already published only cover rectangular co-ordinates. Similar tables for cylindrical and spherical co-ordinates have also been prepared, but are not yet published. Some results for these co-ordinate systems are available elsewhere. For instance, a paper by J. Dougall[14] gives a large number of Green's Functions in cylindrical and spherical co-ordinates, but only for conditions which do not vary with time. Gray, Mathews, and MacRobert[15] refer to this paper, and also work out a number of the results in cylindrical co-ordinates, again for conditions which do not vary with time. When conditions vary with time, if the Green's function for an instantaneous point source of heat can be found in the literature or otherwise, then the Green's function for an instantaneous point charge may be obtained from it by changing the time function T_1^* in the heat problem to T_2^* in the charge problem where

$$T_1^*=\exp\{-[(\alpha_p^2+\beta_q^2+\gamma_r^2)(t-t_1)/g_1]\}\times$$
$$H(t,t_1)/g_1\qquad(65)$$

and

$$T_2^*=\frac{\sin[(t-t_1)(\alpha_p^2+\beta_q^2+\gamma_r^2)^{\frac{1}{2}}/g_2^{\frac{1}{2}}]\,H(t,t_1)}{[(\alpha_p^2+\beta_q^2+\gamma r^2)g_2]^{\frac{1}{2}}}\qquad(66)$$

This procedure is explained in a paper by the author[16] where the preceding equations are given as equations 8.18 and 10.2 respectively. Many expressions for Green's functions in the theory of heat due to instantaneous point sources in cylindrical and spherical co-ordinates will be found, for instance in Carslaw,[17] and these may be translated into electromagnetic theory by this method. It should be noted this procedure applies only when the results are given in terms of normal co-ordinates, so that the term containing time does not involve any of the other co-ordinates.

As an example, Carslaw (on p. 197) gives the Green's function for an instantaneous point source of heat at r_1, θ_1, z_1 inside the space bounded by the cylinder $r=a$ and the planes $z=0$, $z=h$, $\theta=0$ and $\theta=\theta_0$, when the walls are at zero temperature

$$v=\frac{8}{ah\theta_0}\sum_{m=1}\sum_{n=1}\sin\gamma z\sin\gamma z_1\sin\delta\theta\times$$
$$\sin\delta\theta_1\frac{J_\delta(\alpha r)J_\delta(\alpha r_1)}{[J_\delta(\alpha a)]^2}T\qquad(67)$$

where

$$T=\exp[-k(\alpha^2+\gamma^2)(t-t_1)]\qquad(68)$$

α is a positive root of $J_\delta(\alpha a)=0$, $\gamma=m\pi/h$ and $\delta=n\pi/\theta_0$ and m and n take the values of 1, 2, 3 . . . . To get the corresponding field due to an instantaneous point charge, change T to T_1 where

$$T_1=\frac{c\sin[c(t-t_1)(\alpha^2+\gamma^2)^{\frac{1}{2}}]}{(\alpha^2+\gamma^2)^{\frac{1}{2}}}\qquad(69)$$

9. Conclusion

The application of the instantaneous point source or the Green's function to the evaluation of electromagnetic fields has been illustrated by several examples. These range from a straight wire of finite or infinite length in free space, where there are no boundaries at a finite distance, to resonant cavities and wave guides, where definite boundary conditions have to be met. It is surprising in this method of approach how many different problems can be evaluated from a single result obtained from the Tables of Green's Functions.

The procedure described here should help fill a gap which has existed between the two approaches which we may call the methods of circuit theory in the calculation of fields, and the methods based on the calculation of fields from the wave equation of Maxwell. In recent years the distinction between the two approaches has been growing less and less, as is evidenced by the general use of Maxwell's equations in connection with resonant cavities and wave guides. The solution of Maxwell's equation corresponding to the instantaneous point source or the Green's function is admirably adapted to giving expressions for the field due to a charged particle at rest or in motion. From this, it is only a short step to build up the solution for a circuit by integration of the individual charges placed around the circuit. How this can be done is illustrated in the paper by several typical examples, and, at the same time the corresponding fields are derived.

It may also be pointed out that the solutions obtained here contain not only the "steady-state" solution, but also terms corresponding to the "transient" solution, which ensure that at a distance r from the source, the field is zero when $r>c(t-t_0)$, so that a wave travelling with velocity c has not reached the point. These transient terms are not included in the usual theory, but doubtless as further work is done, they will prove useful.

References

1. The Derivation of Vector Potential from Tables for Scalar Potential, **J. J. Smith.** *AIEE Transactions*, vol. 71, pt. I, 1952, pp. 169–74.

2. Tables of Green's Functions, Fourier Series, and Impulse Functions for Rectangular Co-ordinate Systems, **J. J. Smith.** *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 22–30.

3. Electromagnetic Waves (book), **Sergei A. Schelkunoff.** D. Van Nostrand Company, Inc., New York, N. Y., 1943, equation 2-5, p. 130.*

4. Electromagnetic Theory (book), **J. A. Stratton.** McGraw-Hill Book Company, Inc., New York, N. Y., 1941, p. 428.*

5. Operational Calculus (book), **B. Van der Pol, H. Bremmer.** Cambridge University Press, Cambridge, England, 1950, first equation, p. 359.

6. Electricity and Magnetism (book), **J. H. Jeans.** Cambridge University Press, Cambridge, England, second edition, 1911, equation 419, paragraph 496, p. 438.

7. **Schelkunoff,** *op. cit.*, equation 2-8, p. 131.

8. **Schelkunoff,** *op. cit.*, equation 2-13, p. 132.

9. **Stratton,** *op. cit.*, equation 1, p. 440

10. Advanced Antenna Theory (book), **S. A. Schelkunoff.** John Wiley and Sons, Inc., New York, N. Y., 1952, equation 18, p. 131.

11. See equations 42 and 43 of reference 1.

12. See reference 3, pp. 388–89.

13. Principles and Applications of Wave Guide Transmission (book), **C. C. Southworth.** D. Van Nostrand Company, Inc., New York, N. Y., 1950, equation 5.5-5, p. 118.

14. The Green's Function, **John Dougall.** *Proceedings*, Edinburgh Mathematical Society, Edinburgh, Scotland, vol. XVIII.

15. Bessel Functions (book), **Gray, Mathews, MacRobert.** Macmillan and Company, Ltd., London, England, 1952, second edition, pp. 101–10.

16. The Extension of the Heaviside Expansion Theorem, **J. J. Smith.** *Journal of the Franklin Institute*, Philadelphia, Pa., vol. 238, no. 4, Oct. 1944, p. 245.

17. Theory of Heat Conduction, **H. S. Carslaw.** Macmillan and Company, Ltd., London, England, 1921, second edition. (See chapter X on the use of Green's functions, in particular; other results will be found elsewhere in this work.)

---

* For comparison with equation 1 of the paper, write $i=\rho u$.

No Discussion

High-Speed Teletypewriter Equipment for the Armed Services

C. E. SCHULTHEISS
MEMBER AIEE

THE constant striving of the Signal Corps Engineering Laboratories and the foresight and ingenuity of an enterprising laboratory have produced a teletypewriter capable of 60 to 100 word-per-minute operation, suitable for the armed forces. This paper outlines the development; discusses the features making this equipment unique and acceptable to the armed forces; and shows the complete line of equipment as used in an integrated communication system.

History

As early as 1926 the Office of the Chief Signal Officer recognized the need for teletypewriter equipment having features required by a mobile tactical force. These features were extreme portability, lightness, resistance to environment, simplicity in maintenance, and, finally, an ability to operate while being transported, as in a plane or jeep.

These features were not available in any commercial printer, and as late as 1940 it was necessary to standardize on a modification of a commercial model in order to have any supply of teletypewriter equipment available for the armed forces. During this time, however, much progress had been made, both at the Signal Corps Engineering Laboratories and in private laboratories, in producing the desired features.

One such laboratory was that directed by E. F. Kleinschmidt. Mr. Kleinschmidt had been continuously active in research and development of teletypewriter equipment since 1895. The efforts of this laboratory were now directed toward making a lighter higher-speed simplified teletypewriter.

In February 1944, Kleinschmidt Laboratories, Inc., exhibited a progressive stop printer to the officials of the Office of the Chief Signal Officer. Further development led to the standardization of Kleinschmidt teletypewriters for the armed forces in 1948.

Features

The evolution of design resulted in the production of the light-weight printer shown in Fig. 1. This is a completely tactical unit weighing 45 pounds, including a submersible cover (not shown). Fig. 2 is a view with the dust cover removed, showing the simplicity, the reduced number of parts, and the extreme accessibility. Because of simplicity and

---

Paper 54-114, recommended by the AIEE Telegraph Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing December 2, 1953.

C. E. Schultheiss is with Kleinschmidt Laboratories, Inc., Deerfield, Ill.

Fig. 1. TT-4A shown with dust cover

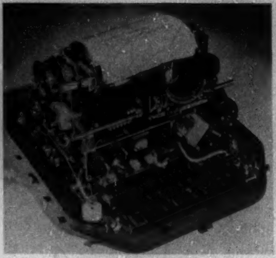


Fig. 2. TT-4A shown with dust cover removed

fewer parts, adjustments and maintenance have been greatly reduced.

Two other operating features are high-speed printing at a lower revolution-per-minute rate, and ease of speed control. A complete explanation of each will be presented under the sections entitled "General Functioning" and "Speed Control."

Unitized Design

In addition to the operating features mentioned, the design of the page printer is such that the separate functions are actually accomplished in physically separate units. As an example, the selector mechanism, shown in Fig. 3, consisting of the selector magnet (removed in this view), selector levers, cam, transfer mechanism, etc., can operate and perform its function separate from other mechanisms. Further, this mechanism can operate in other positions, as it does in a reperforator, see Fig. 14. It was therefore possible to use standard parts throughout the entire line of equipment. The same is true of the code-bar cage, the keyboard, the printing mechanism, the stop-start clutch, the tape-handling mechanism, and many other smaller parts.

Basic Teletypewriter System

To better understand the detailed functioning, it is well to review the basic system and principle of operation. Fig. 4 shows the basic circuit for two interconnected teletypewriters. The sending contacts of either machine will operate both selector magnets in accordance with the signaling code assigned to the character being sent.

The signal code, known as the Baudot code, is made up of five equal length units or elements. Each element may be either a marking or spacing pulse. Fig. 5 is a sample complete signal for a typical letter, in this case, *X*. The two conditions, either marking or spacing, of each element of the code are represented electrically by current or no current in neutral operation, or by the polarity of the current in polar operation. These two conditions may also be represented mechanically by the relationship of five bars, as shown in Fig. 6. The basic operation requires the setting up of the signal code mechanically by code-bar positions; then transforming this position relationship to an electrical and time relationship for transmission; and finally back to a mechanical or position relationship for printing or perforating.

The sending code bars may be positioned by depressing a key on the keyboard. The position of these code bars is transformed into electric sequential pulses (square waves in the ideal signal) by a motor-driven cam. This cam, by means of sensing levers, converts the position of each code bar into closed or open contacts, thus sending a marking or spacing pulse for each element as the five code bars are sensed in turn. The receiving code bars are positioned by a motor driven cam, but the sensing, when receiving, is done by the selector magnet whose armature is pulled up on a marking pulse, or released on a spacing pulse.

As the position of the first code bar is sensed and sent as an electrical pulse, the receiver must sense this first element of the code group to position the first code bar of the receiving teletypewriter. This becomes a problem of synchronizing; no matter how closely the sending and receiving motors were adjusted to the same speed, they would in time be separated. It becomes necessary, therefore, to send a synchronizing pulse which is accomplished in all standard teletypewriter communications by sending a start and stop pulse. Thus, there is a complete stoppage of the selecting mechanism between each code group.

The complete signal sent for each letter or selection is as shown in Fig. 5 with the code for "*X*." It consists of the basic 5-unit code, preceded by a start pulse, which is a spacing or no current condition, and followed after the fifth unit by a stop pulse, which is a marking or current condition. The start pulse is the same length as each unit of the code, but the stop pulse is 1.42 times as long to insure time for completion of the receiving functions and a complete stop. The signal per character sent is made up of 7.42 units.

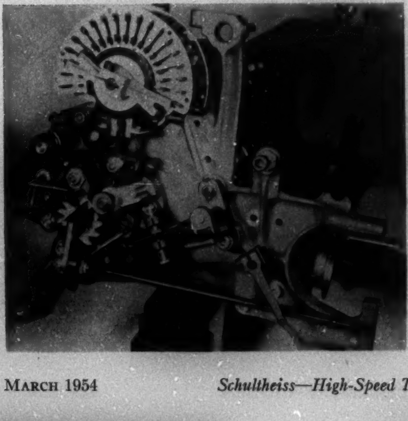


Fig. 3 (left). View of selector mechanism with selector magnet removed

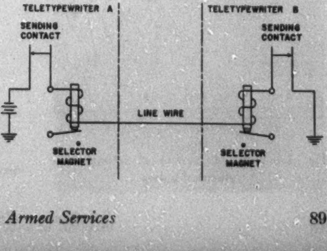


Fig. 4 (right). Basic teletypewriter system, schematic diagram

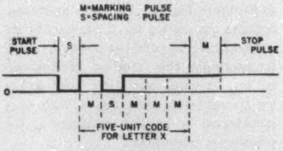


Fig. 5. Standard 5-unit code signal for the letter X

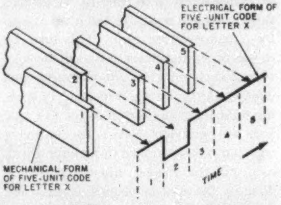
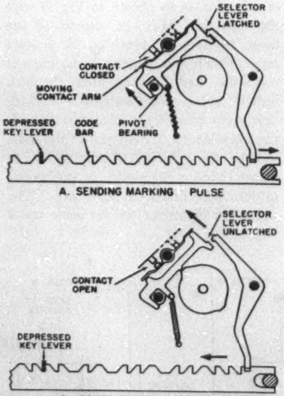

ELECTRICAL FORM OF FIVE-UNIT CODE FOR LETTER X
1
2
3
4
5
MECHANICAL FORM OF FIVE-UNIT CODE FOR LETTER X
1
2
3
4
5
TIME

Fig. 6. Relationship between mechanical (position) and electrical (time) forms of the 5-unit code

General Functioning

This high-speed equipment operates at 60 to 100 words per minute with some units designed for 125 words per minute. To appreciate the design features which allow this high-speed operation, it will be necessary to investigate the timing involved, and the general functioning of a typical printer.

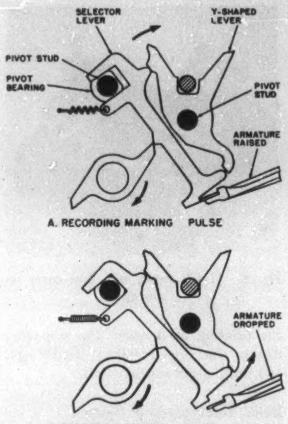
At 100 words per minute there are ten selections per second, allowing a standard of six characters per word. This is a total of 100 milliseconds per code group, which consists of 7.42 units. The duration of the start pulse and each element of the code is 13.5 milliseconds, and for the stop pulse, 19 milliseconds. The speed of operation can best be realized by comparing this to the time of 17 milliseconds for one revolution of the 3,600-rpm motor.


SELECTOR LEVER LATCHED
CONTACT CLOSED
MOVING CONTACT ARM
DEPRESSED KEY LEVER
CODE BAR
PIVOT BEARING
SELECTOR LEVER UNLATCHED
CONTACT OPEN
DEPRESSED KEY LEVER

A. SENDING MARKING PULSE

B. SENDING SPACING PULSE

Fig. 7. Changing mechanical settings into electrical pulses


SELECTOR LEVER
Y-SHAPED LEVER
PIVOT STUD
PIVOT BEARING
PIVOT STUD
ARMATURE RAISED
ARMATURE DROPPED

A. RECORDING MARKING PULSE

B. RECORDING SPACING PULSE

Fig. 8. Pulse-recording portion of the selector mechanism

To accomplish the above speed, the design was conceived to complete an operation in each half revolution, thereby reducing shaft speeds to reasonable values. The reduced shaft speeds mean reduced forces, reduced noise, and reduced wear and maintenance. Also, most cams and stop levers have two working surfaces, one which rests on each operation, thus doubling the useful life of these parts.

The detailed functioning of the sending mechanism is shown in Fig. 7. This outline drawing shows one code bar with its associated sensing and selector lever in each of the marking and spacing positions, *A* and *B* of Fig. 7. The transmitting cam is shown operating, in each case, at roughly the center of a code element. Note that the two lobes on the cam give two operations per revolution. This cam is actually a drum, containing a start lobe, five code lobes, and a stop lobe arranged opposite its associated code bar and stepped around the cam to give a time sequence to the pulses.

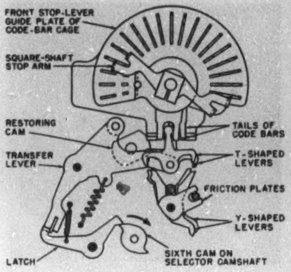

FRONT STOP-LEVER GUIDE PLATE OF CODE-BAR CAGE
SQUARE-SHAFT STOP ARM
RESTORING CAM
TRANSFER LEVER
TAILS OF CODE BARS
T-SHAPED LEVERS
FRICTION PLATES
Y-SHAPED LEVERS
SIXTH CAM ON SELECTOR CAMSHAFT
LATCH

Fig. 9. Transfer mechanism before transfer operation occurs

In the receiving mechanism, outlined in Fig. 8, the sensing lever is replaced by the selector magnet armature. When the selector magnet senses the start pulse, the selector cam is released to revolve. Each section of this compound cam strikes its associated selector lever and drives the *Y*-shaped lever to the right for recording a marking element, Fig. 8(A), or to the left for a spacing element, Fig. 8(B). The selection is thus set up and stored in the mechanical position of the *Y*-shaped levers.

A sixth cam on the selector cam shaft trips the transfer mechanism, Fig. 9, and drops the *T*-shaped levers, which transfer the code as set in the *Y*-shaped

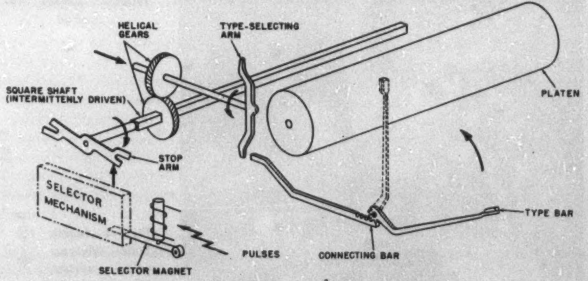

HELICAL GEARS
TYPE-SELECTING ARM
SQUARE SHAFT (INTERMITTENLY DRIVEN)
PLATEN
STOP ARM
SELECTOR MECHANISM
TYPE BAR
PULSES
CONNECTING BAR
SELECTOR MAGNET

Fig. 10. Typing unit, simplified functional diagram

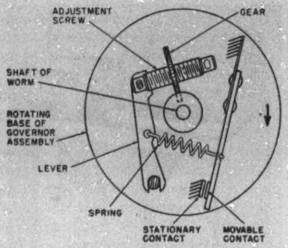


Fig. 11. Governor with target and cover removed

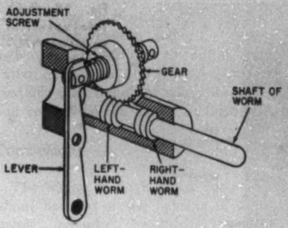


Fig. 12. Worm drive for adjusting governor

levers to positions of the code bars. These bars are so notched that there is one (and only one) position below a stop lever where there will be a notch in each code bar. The transfer action resets the code bars, which cam out the previously selected stop lever and allow the newly selected stop lever to fall in, under the influence of a spring, and intercept the stop arm of the square shaft which was set in motion by the transfer action. Note again the two ends on the stop arm which give two operations per revolution. The position of the square shaft sets the type selecting arm.

Fig. 10 shows how typing is accomplished. The printing bail moves the shaft, holding the type-selecting arm, through the helical gear, striking a hammer blow on the type bar linkage to cause printing. Outstanding features are:

1. The hammer blow of the type bar;
2. The design of the type basket;
3. The arrangement of the type bars;
4. The power-driven carriage;
5. The stationary platen.

Features 1, 2, and 3 combine to allow interference-free operation of the type bars at well over 150 words per minute, giving a comfortable margin over the designed 100-words-per-minute operation.

Feature 4 is very important to a tactical machine and was one reason for selection of this equipment for the armed forces. Being power driven, the carriage will space and return at angles up to 45 degrees. This is a mandatory factor for airborne equipment. Feature 5 minimizes the problem of paper handling, especially fanfold and multiple copy. On all fixed plant equipment the standard platen is a sprocket type to handle either plain paper or multiple copy having perforated margins.

Speed Control

Whenever this equipment is designed as a completely tactical unit, the motor used must be universal, since power supplies in tactical areas are unregulated. The problem is, then, one of speed control. The equipment can tolerate considerable speed variation between sending and receiving units when there is an ideal signal, but any speed change between one synchronizing stop pulse and the following does change the point of selection from the ideal. No signal, of course, is ideal, and all contain some distortion. One measure of the equipment is the amount or margin of distortion it can accept. Speed variations reduce this margin and must therefore be held to a minimum.

The standard for speed control is a 180-vibrations-per-second tuning fork, with shutters through which a 3-dot target is viewed; see Fig. 2.

The motor speed is controlled by the opening or closing of spring-balanced contacts which are mounted as a governor unit directly on the motor shaft. The shaft is controlled at a nominal speed of 3,600 rpm. The contacts are positioned so that they are affected both by centrifugal and acceleration forces. With reference to Fig. 11, the movable contact is seen to be balanced by a spring and lever assembly positioned by an adjusting screw. This adjusting screw forms the axis of a gear which engages a left or right hand worm, see Fig. 12. This whole assembly revolves with the motor shaft. The worm shaft is floating, and may be gripped between the forefinger and thumb and held so either the left or right hand worm engages the gear. The gear then revolves clockwise, as viewed from the governor end, around the now-stationary worm in a manner which drives the adjustment screw to the left or right, resulting in an increase or decrease of motor speed. Speed is easily controlled within plus or minus one half of one percent.

The Communication System

There is available to the armed forces a complete line of high-speed teletypewriter equipment from the tactical lightweight portable page printer to a fully automatic switching center. Each unit makes use of the same basic design of transmitting units, selector mechanism, page-printing mechanism, tape perforating and printing mechanism, and other minor units.

Fig. 13 shows a communication system using all the available units. This figure will be used as a reference to show how

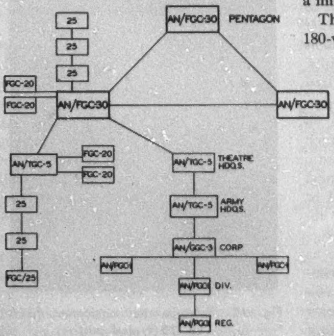


Fig. 13 (left). Schematic of a possible communication system

Fig. 14 (right). Teletypewriter equipment AN/GGC-3

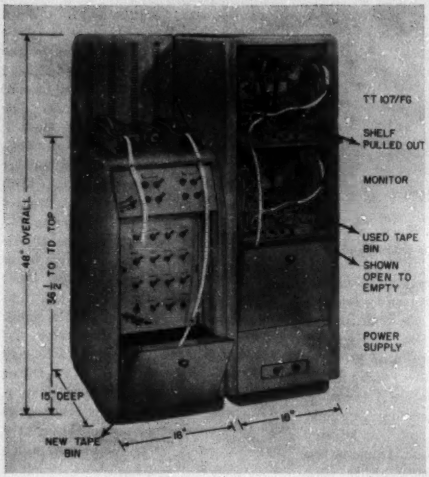


Fig. 15 (left). Teletypewriter equipment AN/-TGC-5

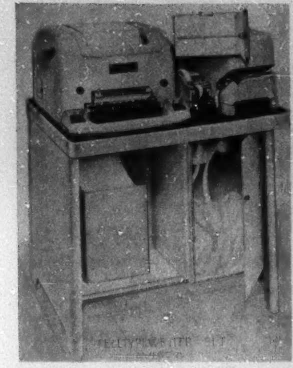


Fig. 17 (right). Teletypewriter equipment AN/-FGC-25

each of the separate equipments fit into the integrated system.

Previous discussion has introduced the *TT4* light-weight page printer, Fig. 1. This printer forms part of a set *AN/-PGC-1*, including carrying case and accessories. It is highly transportable for use by the smaller army organizational units up through the division. At corps headquarters, the reperforator unit *AN/-GGC-3*, as shown in Fig. 14, would be introduced. At this point, communication enters the automatic transmission stage. The reperforator cuts the tape which is now available for manual insertion into the proper tape distributor associated with each reperforator. Routing and starting of transmission are manual. This equipment is considered tactical.

Fig. 16. Teletypewriter equipment AN/-FGC-20

At least one *AN/GGC-3* unit will be associated with the army headquarters, where the semi-automatic or torn-tape switching center, *AN/TGC-5*, is introduced. A typical unit is shown in Fig. 15. A theater headquarters may be similarly equipped. This equipment is transportable and may be expanded to any number of lines desired.

It is understood, of course, that any number of interconnections to other headquarters or automatic switching centers may be made. Likewise, there would be associated with each *AN/TGC-5* center outlets for local delivery or connections for incoming messages from smaller units.

Local deliveries are made on the fixed plant page printers *AN/FGC-20*, shown in Fig. 16. Each is equipped with its own table and power supply and has such refinements as quiet operation, automatic carriage return and line feed, manual reset to "letters," use of either roll or fanfold paper, and, finally, a paper rewind as optional equipment.

Where messages are to originate, as well as be delivered, as in a switching center or smaller fixed positions, an *AN/-FGC-25* unit is used. This unit is shown in Fig. 17. These units may be used on a loop, with each having its own call. The *FGC-25* unit has a page printer, perforator, transmitter-distributor, and a reperforator of special design so that a selector mechanism is continually monitoring the incoming line. When a proper call signal is received, the selection is automatically transferred to the reperforator or page printer, as the operator may choose, and the message recorded. Mes-

Fig. 18. Teletypewriter equipment AN/-FGC-30 (typical unit)

sages can be made up or relayed at this point through the transmitter-distributor, which is part of the set.

The *AN/FGC-30* central office is a fully automatic switching center in which messages with several degrees of priority and with multiple addresses are processed, along with simple through messages. Use is made of cross-office units, which must be run at a higher speed (up to 125 wpm), to prevent blocking an incoming line. Throughout, teletypewriter equipment in the switching center is an adaptation of the basic unitized designs. A typical unit is shown in Fig. 18. Besides the basic units, automatic operation requires other special units such as tape storage indicators, tape motion alarms, tape readers, and other minor indicators. In addition to the teletypewriter equipment, the switching center contains the "brain" of relays directing and controlling the numerous functions.

Conclusions

The stress on high speed has come about because of the benefits of automatic message handling. Teletypewriters operating at 100 words per minute no longer limit the operator. Some units are presently designed for 125-word-prominute operation, and there are pessibilities of even greater speeds.

Automatic operation, however, pushes the limits of equipment. The inertia phenomena of moving levers and of the electrical circuit become increasingly important as a limit is reached. Transferring a message to the printed page by simple mechanical means has a limit since type must be moved in and out of position. The future for teletypewriter equipment lies in the production of high-speed light-weight simple economical units for use at origin and delivery points of messages.

◆

No Discussion

Analysis of Linear Time-Varying Circuits by the Brillouin-Wentzel-Kramers Method

LOUIS A. PIPES
MEMBER AIEE

Synopsis: This paper presents a general mathematical technique for studying the response of linear time-varying circuits that contain parameters whose magnitudes vary in a periodic manner with the time. The method presented is based on the Brillouin-Wentzel-Kramers (BWK) procedure, which has been widely used in connection with problems of wave mechanics. This method gives approximate results with very little error if the variable circuit elements exhibit small variations about a large average value. The application of the method to circuit problems is illustrated by using it to study the free and forced oscillations of typical series circuits that contain periodic capacitance, resistance, and inductance parameters.

THE general behavior of electric circuits whose parameters are constants has been extensively studied and is well understood. In recent years, a great deal of attention has been paid to the general theory and performance of circuits whose parameters are functions of the time, especially when they vary periodically.[1-9] Examples of linear time-varying circuits of practical importance occur in the theory of electrical communications. Frequency modulation, for instance, utilizes variations of capacitance or, to a lesser degree, inductance. The microphone transmitter contains a variable resistance whose value is varied by some source of energy outside the circuit, and the capacitor microphone contains a variable capacitance. The introduction of superregeneration has especially stimulated an interest in circuits that contain periodic variation of a resistance parameter.[10]

The mathematical analysis of the performance of electric circuits whose parameters are periodic functions of the time leads generally to the solution of Mathieu's and Hill's differential equations. These equations have the following general form

$$\frac{d^2y}{dx^2}+G^2(x)y=H(x) \qquad (1)$$

where $G^2(x)$ is a periodic function of x and $H(x)$ is a continuous function of x. Despite the vast amount of literature on the theory of Mathieu functions,[11] it is still a formidable task to obtain rigorous solutions of equation 1 in cases of practical importance, and approximate methods must be employed.

BWK Approximation

The great majority of linear time-varying circuits that occur in communications engineering have the property of the variable parameters involved exhibiting only small variations about a large average value. In such cases a very good approximate solution of equation 1 can be obtained by a method used by Brillouin, Wentzel, and Kramers to solve equations of the type

$$\frac{d^2y}{dx^2}+G^2(x)y=0 \qquad (2)$$

in connection with problems of wave mechanics.[12,13] Although the BWK method was used by Brillouin, Wentzel, and Kramers to solve problems in wave mechanics in 1926, the procedure is a very old one and the first indication of its use is found in the collected papers of the eminent mathematician Liouville, published in 1837.

To determine the type of approximation involved in the BWK procedure, consider the following function

$$y=\frac{1}{\sqrt{G(x)}}\left[C_1\epsilon^{j\phi(x)}+C_2\epsilon^{-j\phi(x)}\right] \qquad (3)$$

where ϵ is the base of the natural logarithms, C_1 and C_2 are arbitrary constants, and $j=\sqrt{-1}$. The function $\phi(x)$ is given by

$$\phi(x)=\int G(x)dx \qquad (4)$$

By direct differentiation of equation 3 it can be shown that this function satisfies the following differential equation

$$\frac{d^2y}{dx^2}+\left[G_2+\frac{G''}{2G}-\frac{3}{4}(G'/G)^2\right]y=0 \qquad (5)$$

If $G^2(x)$ is a periodic function that

Paper 54-183, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted October 19, 1953; made available for printing December 28, 1953.

Louis A. Pipes is with the University of California, Los Angeles, Calif.

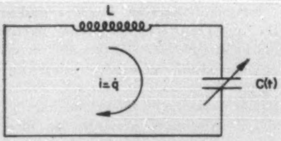


Fig. 1. Capacitance modulation

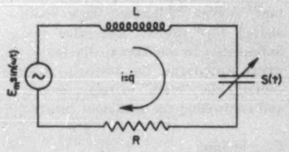


Fig. 2. Circuit with variable elastance

exhibits only small variations about a large average value it can be shown that

$$|G^2(x)|>>|[G''/2G-3(G'/2G)^2]| \qquad (6)$$

throughout the range in x, and hence that equation 3 is an approximate solution of equation 2.

In the case where $G^2(x)$ is a positive periodic function of x, equation 3 may be written in the alternative form

$$y=\frac{1}{\sqrt{G(x)}}\{A\cos[\phi(x)]+B\sin[\phi(x)]\} \qquad (7)$$

where A and B are arbitrary constants, and $\phi(x)$ is given by 4. Equation 7 is the usual BWK approximate solution of equation 2.

Capacitance Modulation

As a first example of the use of the general BWK procedure, consider the circuit of Fig. 1. This is a series circuit that consists of a constant inductance L in series with a capacitance $C(t)$ that varies with time in such a manner that

$$C(t)=C_0[1+2h\cos(2pt)] \qquad (8)$$

where C_0 and h are constant parameters with $h^2<<1$. If q denotes the charge on the capacitance, the differential equation of the circuit is

$$\frac{d^2q}{dt^2}+\frac{q}{LC_0[1+2h\cos(2pt)]}=0 \qquad (9)$$

Typical values of the various parameters in the case of radio frequency modulation are

$$\omega_0-1/\sqrt{LC_0}=2\pi(5\times10^7)\text{ seconds}^{-1} \qquad (10)$$

$$2p=2\pi(5\times10^3)\text{ seconds}^{-1} \qquad (11)$$

$$h=2\times10^{-4} \qquad (12)$$

It is therefore justifiable to neglect terms containing h^2 and higher, and write

$$\frac{1}{[1+2h\cos(2pt)]}=[1-2h\cos(2pt)] \qquad (13)$$

Consequently equation 9 may be written in the following form

$$\frac{d^2q}{dt^2}+\omega_0^2[1-2h\cos(2pt)]q=0 \qquad (14)$$

If the new variables, $y=q$ and $x=pt$, are introduced, equation 14 takes the following form

$$\frac{d^2y}{dx^2}+(\omega_0/p)^2(1-2h\cos 2x)y=0 \qquad (15)$$

This equation is a special case of equation 2 with

$$G(x)=(\omega_0/p)\sqrt{(1-2h\cos 2x)}$$
$$=(\omega_0/p)(1-h\cos 2x) \qquad (16)$$

provided that terms of order h^2 and higher are neglected. Since $(\omega_0/p)=2\times10^4$ and $h=2\times10^{-4}$, it can be seen that equation 6 is satisfied, and, an accurate approximate solution of equation 15 is given by equation 7. In this case we have

$$\phi(x)=\int G(x)dx=(\omega_0/p)\left(x-\frac{h}{2}\sin 2x\right) \qquad (17)$$

Therefore the BWK solution of equation 15 is

$$y=[(\omega_0/p)(1-h\cos 2x)]^{-1/2}(A\cos[ax-$$
$$c\sin 2x]+B\sin[ax-c\sin 2x]) \qquad (18)$$

where A and B are arbitrary constants and $a=(\omega_0/p)$ and $c=\omega_0h/2p$.

To determine the various harmonic components, the following expansions involving Bessel functions[14] may be used

$$\cos(a\sin b)=J_0(a)+2[J_2(a)\cos(2b)+$$
$$J_4(a)\cos(4b)+\ldots]\times$$
$$\sin(a\sin b)=2[J_1(a)\sin(b)+$$
$$J_3(a)\sin(3b)+\ldots] \qquad (19)$$

where the $J_n(a)$ functions are Bessel functions of the first kind and the nth order. With the aid of equation 19, equation 18 may be written in the following form after certain algebraic reductions

$$y=q=K[(\omega_0/p)(1-h\cos 2pt)]^{-1/2}$$
$$\sum_{n=-\infty}^{+\infty}J_n(\omega_0h/2p)\cos[\omega_0-2np)t-\theta] \qquad (20)$$

where K and θ are arbitrary constants. This is the BWK solution of equation 14 for the circuit charge. If the quantity h in the square root term of equation 20 is neglected in comparison with unity, this equation may be written in the form

$$q=K_0\sum_{n=-\infty}^{+\infty}J_n(\omega_0h/2p)\cos[(\omega_0-2np)t-\theta] \qquad (21)$$

where K_0 and θ are arbitrary constants. Equation 21 is a less accurate solution of equation 13 and was obtained by Carson[15] by a different procedure.

Forced Oscillations of a Circuit with Variable Elastance

As a second example of the use of the BWK procedure, consider the circuit of Fig. 2. This circuit consists of a series connection of a constant resistance R a constant inductance L and a time-varying elastance $S(t)$ in series with a periodic electromotive force $E(t)=E_m\sin(\omega t)$. Let it be assumed that the variation of the elastance is of the form

$$S(t)=S_0[1-2h\cos(2pt)]\text{ where }h^2<<1 \qquad (22)$$

S_0 is the constant part of the elastance and h is a small amplitude parameter. The differential equation satisfied by the charge q on the elastance is

$$\frac{Ld^2q}{dt^2}+\frac{Rdq}{dt}+S_0[1-2h\cos(2pt)]q$$
$$=E_m\sin(\omega t) \qquad (23)$$

Let

$$b=R/2L,\ \omega_0=\sqrt{S_0/L} \qquad (24)$$

and introduce the new variable $y(t)$ by means of the transformation

$$q(t)=\epsilon^{-bt}y(t) \qquad (25)$$

In terms of these parameters, equation 23 may be written in the form

$$\frac{d^2y}{dt^2}+[(\omega_0^2-b^2)-2h\omega_0^2\cos(2pt)]y$$
$$=\frac{E_m}{L}\epsilon^{bt}\sin(\omega t) \qquad (26)$$

Let $x=pt$, then equation 26 becomes

$$\frac{d^2y}{dx^2}+\frac{1}{p^2}[(\omega_0^2-b^2)-2h\omega_0^2\cos(2x)]y$$
$$=\frac{E_m}{Lp}\epsilon^{bx/p}\sin(\omega x/p) \qquad (27)$$

If the resistance R of the circuit is small, then $b^2<<\omega_0^2$ and the terms b^2 may be neglected in equation 27. The complementary function of the differential equation 27 is the solution of the equation

$$\frac{d^2y}{dx^2}+(\omega_0^2/p^2)[1-2h\cos(2x)]y=0 \qquad (28)$$

This is equation 16 and its BWK solution is given by equation 18. If h is small, it can be neglected in comparison with unity in the square-root term of equation 18, and the general solution of equation 28 may be written in the form

$$y(x)=A_1y_1(x)+A_2y_2(x) \qquad (29)$$

where A_1 and A_2 are arbitrary constants and functions $y_1(x)$ and $y_2(x)$ are given by

$$y_1(x)=\cos\,[ax-c\sin(2x)]$$

$$y_2(x)=\sin\,[ax-c\sin(2x)],\ a=(\omega_0/p),\quad c=\omega_0h/2p \qquad (30)$$

The general solution of equation 27 for the forced oscillations of the circuit of Fig. 2 may now be obtained by the method of the variation of parameters,[16] in the following form

$$y(x)=A_1y_1(x)+A_2y_2(x)+y_2(x)\times \int\frac{R}{W}(x)y_1(x)dx-y_1(x)\int R(x)y_2(x)dx \qquad (31)$$

where $R(x)$ is the right member of equation 27 given by

$$R(x)=(E_m/Lp^2)\epsilon^{bx}/p\sin(\omega x/p) \qquad (32)$$

and W is the Wronskian of the solutions, equation 30, whose value is

$$W=(y_1y_2'-y_2y_1')=\omega_0/p=a \qquad (33)$$

for the case in which $h<<1$. The terms containing the arbitrary constants in equation 31 constitute the complementary function of equation 27, and the terms involving the indefinite integrals give the particular integral of the equation. The approximate general solution of equation 23 may now be obtained by expressing equations 25 and 33 in terms of the original variables. It has the following form

$$\begin{aligned}q(t)={}&A_1\epsilon^{-bt}\cos\,[\omega_0t-c\sin(2pt)]+\\&A_2\epsilon^{-bt}\sin\,[\omega_0t-c\sin(2pt)]+\\&(E_m/\omega_0L)\epsilon^{-bt}\sin\,[\omega_0t-c\sin(2pt)]\times\\&\int\epsilon^{bt}\sin(\omega t)\cos\,[\omega_0t-c\sin(2pt)]dt-\\&(E_m/\omega_0L)\epsilon^{-bt}\cos\,[\omega_0t-c\sin(2pt)]\times\\&\int\epsilon^{bt}\sin(\omega t)\sin\,[\omega_0t-c\sin(2pt)]dt\end{aligned} \qquad (34)$$

The arbitrary constants A_1 and A_2 of the transient part of the solution may be evaluated if the initial conditions of the circuit at $t=0$ are given. The harmonic content of the steady-state solution may be obtained by expanding the integrand as a series of harmonic terms by means of equation 19 and integrating term by term.

Series Circuit with Periodically Varying Resistance

In recent years, the introduction of superregeneration has stimulated an interest in the study of circuits that involve a periodic variation of the resistance parameter.[7] Consider the circuit of Fig. 3. This circuit consists of a constant inductance and capacitance in series with a periodically varying resistance, R(t). It is assumed that a harmonic forcing potential $E(t)=E_m\sin(\omega t)$ is impressed on the circuit.

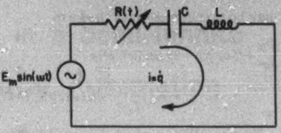


Fig. 3. Circuit with variable resistance

The circulating charge of this circuit satisfies the following differential equation

$$\frac{d^2q}{dt^2}+\frac{R}{L}(t)\frac{dq}{dt}+\frac{q}{LC}=\frac{E_m}{L}\sin(\omega t) \qquad (35)$$

Let the resistance parameter have the following periodic variation

$$R(t)=R_0[1+k\cos(pt)]\ \text{where}\ k^2<<1 \qquad (36)$$

Let

$$u(t)=\frac{1}{2L}\int R(t)dt=b\left[t+\frac{k}{p}\sin(pt)\right] \qquad (37)$$

with

$$b=R_0/2L \qquad (38)$$

and introduce the transformation

$$q(t)=\epsilon^{-u(t)}y(t) \qquad (39)$$

This transformation transforms equation 35 into the following equation

$$\frac{d^2y}{dt^2}+[1/LC-R^2/4L^2-R'/2L]y=\frac{E_m}{L}\epsilon^{u(t)}\sin(\omega t) \qquad (40)$$

If equation 36 is substituted into equation 40 and terms of the order k^2 are neglected, the result is

$$\frac{d^2y}{dt^2}+[\omega_0^2+kb(p\sin pt-2b\cos pt)]y=\frac{E_m}{L}\epsilon^{u(t)}\sin(\omega t) \qquad (41)$$

where

$$\omega_0=\sqrt{(1/LC-b^2)} \qquad (42)$$

If b is small, $p>>2b$ and the cos (pt) term in equation 41 may be neglected. The complementary function of equation 41 is then the solution of

$$\frac{d^2y}{dt^2}+\omega_0^2[1+(kbp/\omega_0^2)\sin(pt)]y=0 \qquad (43)$$

Let $x=pt$. In the variable x, equation 43 is transformed into

$$\frac{d^2y}{dx^2}+(\omega_0/p)^2[1+(kbp/\omega_0^2)\sin(x)]y=0 \qquad (44)$$

The function $G(x)$ (of the section on BWK approximation) is now given by

$$G(x)=(\omega_0/p)\sqrt{[1+(kbp/\omega_0^2)\sin(x)}=[(\omega_0/p+m\sin(x] \qquad (45)$$

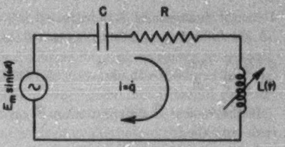


Fig. 4. Circuit with variable inductance

where

$$m=kb/2\omega_0 \qquad (46)$$

The function $\phi(x)$ is in this case given by

$$\phi(x)=\int G(x)dx=[\omega_0x/p-m\cos(x)] \qquad (47)$$

Two linearly independent BWK solutions of equation 44 are

$$\begin{aligned}y_1(x)&=\cos\,[\omega_0x/p-m\cos(x)]\\y_2(x)&=\sin\,[\omega_0x/p-m\cos(x)]\end{aligned} \qquad (48)$$

By the method described in the section "Forced Oscillations of a Circuit with Variable Elastance," and after some algebraic reductions, the following approximate general solution of equation 35 may be obtained

$$\begin{aligned}q(t)={}&A_1\epsilon^{-u(t)}\cos\,[\omega_0t-m\cos(pt)]+\\&A_2\epsilon^{-u(t)}\sin\,[\omega_0t-m\cos(pt)]+\\&(E_m/\omega_0L)\epsilon^{-u(t)}\sin\,[\omega_0t-m\cos(pt)]\times\\&\int\epsilon^{u(t)}\sin(\omega t)\cos\,[\omega_0t-m\cos(pt)]\times\\&dt-(E_m/\omega_0L)\epsilon^{-u(t)}\times\\&\cos\,[\omega_0t-m\cos(pt)]\int\epsilon^{u(t)}\sin(\omega t)\times\\&\sin\,[\omega_0t-m\cos(pt)]dt\end{aligned} \qquad (49)$$

The quantities A_1 and A_2 are the arbitrary constants of the transient part of the solution and the terms involving the integrals give the steady-state response of the system to the BWK order of approximation.

Series Circuit with Periodically Varying Inductance

As a final example of the BWK method, consider the circuit of Fig. 4. This circuit consists of a constant resistance and capacitance in series with a periodically varying inductance parameter. The circuit is energized by a periodic electromotive force $E(t)=E_m\sin(\omega t)$.

The differential equation satisfied by the charge on the capacitance is[8]

$$\frac{d\phi}{dt}+\frac{Rdq}{dt}+\frac{q}{C}=E_m\sin(\omega t) \qquad (50)$$

where ϕ is the total magnetic flux linking the circuit, R is the resistance, and C the capacitance of the circuit. The flux ϕ is related to the inductance of the circuit by the fundamental equation

$$\phi=Li=\frac{Ldq}{dt} \qquad (51)$$

Hence if equation 51 is substituted into 50, the following equation is obtained

$$\frac{d}{dt}\left(\frac{Ldq}{dt}\right)+\frac{Rdq}{dt}+\frac{q}{C}=E_m \sin(\omega t) \qquad (52)$$

If the indicated differentiation is carried out, then

$$\frac{Ld^2q}{dt^2}+\left(R+\frac{dL}{dt}\right)q+\frac{q}{C}=E_m \sin(\omega t) \qquad (53)$$

This equation may be written in the following form

$$\frac{d^2q}{dt^2}+2P(t)\frac{dq}{dt}+\frac{q}{LC}=\frac{E^m}{L}\sin(\omega t) \qquad (54)$$

where

$$P(t)=\frac{1}{2L}\left(R+\frac{dL}{dt}\right) \qquad (55)$$

Let

$$v(t)=\int P(t)dt \qquad (56)$$

$$q(t)=\epsilon^{-v(t)}y(t) \qquad (57)$$

then equation 54 is transformed into

$$\frac{d^2y}{dt^2}+\left[1/LC-\frac{dP}{dt}-P^2\right]y =\frac{E_m}{L}\epsilon^{v(t)}\sin(\omega t) \qquad (58)$$

Let it now be assumed that the variable inductance has the following form

$$L(t)=L_0[1+a\sin(pt)], \text{ where } a^2<<1 \qquad (59)$$

If equation 59 is substituted into equation 55, 56, and 58, and terms in a^2 are neglected, then after some algebraic reductions equation 58 may be written in the following form

$$\frac{d^2y}{dt^2}+[\omega_0^2+2g\sin(pt)]y=\frac{E_m}{L}\epsilon^{v(t)}\sin(\omega t) \qquad (60)$$

where

$$\omega_0=\sqrt{(1/L_0C-R^2/4L_0^2)} \qquad (61)$$

$$g=\frac{a}{4}(p^2-2/L_0C+R^2/L_0^2) \qquad (62)$$

$$v(t)=\int P(t)dt=\frac{Rt}{2L_0}+\frac{a}{2}\times \left[\sin(pt)+\frac{R}{L_0p}\cos(pt)\right] \qquad (63)$$

By the transformation $x=pt$ equation 60 may be transformed into the form

$$\frac{d^2y}{dx^2}+[(\omega_0/p)^2+2g\sin(x)/p^2]y =E_m\frac{\epsilon^{v(x/p)}}{Lp^2}\sin(\omega x/p) \qquad (64)$$

By the BWK method discussed under "BWK Approximation," the following complementary function of equation 64 y_c may be obtained

$$y_c=A_1\cos[\omega_0x/p-z\cos(x)]+ A_2\sin[\omega_0x/p-z\cos(x)] \qquad (65)$$

where

$$z=g/\omega_0p \qquad (66)$$

A_1 and A_2 in equation 65 are arbitrary constants. The method of the variation of parameters, discussed in the section on forced oscillations, gives the following approximate general solution of equation 53 with the inductance variation equation 59

$$\begin{aligned}q(t)=A_1\epsilon^{-v(t)}\cos[\omega_0t-z\cos(pt)]+ \\ A_2\sin[\omega_0t-z\cos(pt)]\epsilon^{-v(t)}+ \\ (E_m/\omega_0L)\epsilon^{-v(t)}\sin[\omega_0t-z\cos(pt)]\times \\ \int\epsilon^{v(t)}\sin(\omega t)\cos[\omega_0t-z\cos(pt)]- \\ (E_m/\omega_0L)\epsilon^{-v(t)}\sin[\omega_0t-z\cos(pt)]\times \\ \int\epsilon^{v(t)}\sin(\omega t)\cos[\omega_0t-z\cos(pt)]\end{aligned} \qquad (67)$$

The transient part of the solution contains the arbitrary constants and the integral terms contain the steady-state response. The harmonic content of the steady-state response may be obtained by using the Bessel function expansions, equation 19, and term by term integration.

Conclusion

The analysis presented in this paper is an attempt to apply the method of the BWK approximation, which has yielded very useful and interesting results in problems of wave mechanics to the analysis of linear circuits whose parameters vary periodically with the time. A survey of the literature of circuit analysis reveals that this powerful method of analysis has been neglected by the electrical engineers, and it is hoped that this presentation will give an impetus to its future use for the solution of problems involving linear, time-varying parameters.

References

1. Operational and Matrix Methods in Linear Variable Networks, **L. A. Pipes.** *The Philosophical Magazine*, London, England, vol. 25, 1938, pp. 585–600.

2. On the Oscillations of a Circuit Having a Periodically Varying Capacitance, **W. L. Barrow.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 22, 1934, pp. 201–12.

3. Ueber die freien Schwingungen in Kondensatorkreisen mit periodisch veralnderlicher Kapazitaet, **A. Erdelyi.** *Physica*, Utrecht, Netherlands, vol. 5, no. 19, 1934, pp. 585–622.

4. Solution of Variable Circuits by Matrices, **L. A. Pipes.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 224, 1937, pp. 767–77.

5. A General Review of Linear Varying Parameter and Nonlinear Circuit Analysis, **W. R. Bennett.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 38, 1950, pp. 259–63.

6. Resonant Circuits with Time-Varying Parameters, **R. H. Kingston.** *Proceedings*, Institute of Radio Engineers, vol. 37, 1949, pp. 1478–81.

7. Resonant Circuit with Periodically-Varying Parameters, **P. Bura, D. M. Tombs.** *Wireless Engineer*, London, England, April 1952, pp. 95–100; May 1952, pp. 120–26.

8. An Electrical Network with Varying Parameters, **C. P. Gadsen.** *Quarterly of Applied Mathematics*, Brown University, Providence, R. I., vol. 8, no. 2, 1950, pp. 199–205.

9. Transforms for Linear Time-Varying Systems, **J. A. Aseltine.** *Report No. 52.1*, University of California, Los Angeles, Calif., Feb. 1952.

10. Vacuum-Tube Oscillators (book), **W. A. Edson.** John Wiley and Sons, Inc., New York, N. Y., 1953, pp. 408–12.

11. Theory and Application of Mathieu Functions (book), **N. W. McLachlan.** Oxford University Press, London, 1947. (Contains 226 references to the international literature.)

12. The Fundamental Principles of Quantum Mechanics, **E. C. Kemble.** McGraw-Hill Book Company, Inc., New York, N. Y., 1937, pp. 90–112. (Contains a detailed exposition of the BWK procedure and an extensive list of references to papers on wave mechanics.)

13. The B.W.K. Approximation and Hill's Equation, **L. Brillouin.** *Quarterly of Applied Mathematics*, Brown University, Providence, R. I., vol. 7, no. 4, 1950, pp. 363–80.

14. Bessel Functions for Engineers (book), **N. W. McLachlan.** *Oxford University Press*, London, England, 1934, pp. 42–43.

15. Notes on the Theory of Frequency Modulation, **J. R. Carson.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 10, no. 62, 1922.

16. Differential Equations (book), **L. R. Ford.** McGraw-Hill Book Company, Inc., New York, N. Y., 1933, pp. 75–76.

◆

No Discussion

Erratum

In the paper entitled "An Analysis of an Analogue Solution Applied to the Heat Conduction Problem in a Cartridge Fuse" by A. E. Guile and E. B. Carne which appeared in the January 1954 issue of Communication and Electronics, pages 861–68, the key to Fig. 6 was omitted. The key should be:

——— Theoretical solution, no radial loss

— — — — Analogue solution, no radial loss

. . . . Analogue solution, with radial loss

Eddy-Current Losses in a Semi-Infinite Solid due to a Nearby Alternating Current

H. PORITSKY
NONMEMBER AIEE

R. P. JERRARD
NONMEMBER AIEE

IN many engineering structures one is confronted with the problem of eddy-current losses in conducting solids due to nearby alternating currents. In electric machines such as motors, generators, and transformers, the machine parts exposed to variable fields are made up of laminations so as to break up the eddy-current paths and avoid losses. Yet, in parts of these machines, solid metallic structures do occur which are exposed to alternating fields and which it is impossible to laminate. Among these are end plates of rotors and stators in the field of the end turns, and transformer tanks in the field of current leads. Similarly, in buildings of reinforced steel construction, heavy current leads may cause eddy-current losses in the steel girders. Finally, eddy currents may be excited in sea water or in the earth, either under conditions of zero-phase short-circuit currents in power lines or when alternating currents are provided with ground leads. In problems of interference between power and telephone lines, the distribution of eddy currents and their effect on the interference factors are of interest.

In view of the above, it is important to have available typical solutions for eddy currents and losses due to them in extended structures. One such structure is a semi-infinite conducting permeable solid bounded by a plane, and exposed to the field of an alternating current flowing parallel to it, as shown in Fig. 1. The electromagnetic field for this case, the exact distribution of the eddy currents, and the losses due to them are studied in the paper. In the case of the earth or sea water, the approximation by a semi-infinite solid is very good indeed. In the case of machine structures or steel structures, no infinite solid actually occurs; yet a steel plate whose thickness is greater than the depth of the penetration at the frequency in question (presently to be defined) will act essentially like a semi-infinite solid.

The field determination is carried out by solving Maxwell's equations both in air space and in the conducting solid, and their solution yields both the induced electric field and the magnetic field at every point. The solution is carried out for an arbitrary value of conductivity λ, permeability μ of the conducting solid, distance h of the primary or exciting current from the plane boundary, and frequency of this current. The permeability μ is assumed to be constant, so that no effects of saturation are taken into account. Curves are provided in dimensionless form from which the complete losses due to the eddy currents can be found. The results agree with those obtained by J. R. Carson[1] for $\mu=1$, and by N. H. Wise[2] for values of μ close to 1.

The procedure employed in solving Maxwell's equations is that of Fourier-integral resolution. First the field due to a sinusoidal current sheet at a height above the surface of the solid is calculated. By using proper Fourier-integral superposition corresponding to different wave lengths of the current sheet, one obtains the field due to a single concentrated current. The resulting expression for the field then turns out to be a definite integral.

The sinusoidal-current-sheet components used in the above superposition are not without interest in themselves, since pulsating or traveling sinusoidal current sheets are of frequent occurrence in electrical machinery. By using other Fourier-integral superpositions, fields and eddy currents due to certain distributed exciting currents are obtained, but these are not evaluated in detail.

In addition to the results just noted for the equivalent current impedance, expressions are obtained for the actual field. In particular, the induced electric field **E** parallel to the current is obtained as a definite integral, as shown in section 4. While no tables are obtained for this integral due to the fact it would depend upon too many parameters, complete methods of evaluation of this integral are presented. By following these methods one can then calculate not only the induced field at points other than the primary current location, but one may also calculate the intensity of the eddy currents at any place inside the conducting solid. Thus, if there are several currents, for instance three 3-phase parallel currents, the net effect on the impedance of each one can be obtained by superposing the induced fields due to each current separately; the latter can be calculated from the integrals just mentioned.

A further result of interest will now be mentioned. The question is often asked, what image under the surface of a solid will produce the same effect as is actually produced by the solid itself? Thus, in the study of voltage surges on transmission lines, L. V. Bewley[3] utilized an electrostatic image for a transmission line, as well as an electromagnetic one. The former lay at a depth below the ground equal to the height of the transmission line above it. The latter lay at a different depth and presumably produced the same flux linkage at the transmission line as was actually produced by the distributed eddy currents underground. In the following, the question of possible images of these alternating currents is discussed (though only the steady-state and not the transient problem is considered). It is shown that a single image current is insufficient to produce an induced electric field which is everywhere the same as results from the superposition of the field of the primary and all the eddy currents; a proper distribution of current images lying on the same vertical plane as the primary current is required. Evaluation of this continuous-current image pattern is discussed.

1. Results

Inside the conducting medium, for frequencies for which displacement currents

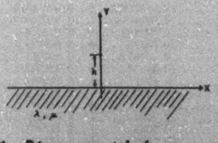


Fig. 1. Primary current J, frequency ω/2π, near semi-infinite solid

---

Paper 54-106, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted October 21, 1953; made available for printing November 27, 1953.

H. Poritsky and R. P. Jerrard are with the General Electric Company, General Engineering Laboratory, Schenectady, N. Y.

The authors gratefully acknowledge the assistance of Dr. G. Horvay, Dr. O. J. Farrell, and A. R. Clairmont in the reduction of the integrals and the numerical calculations.

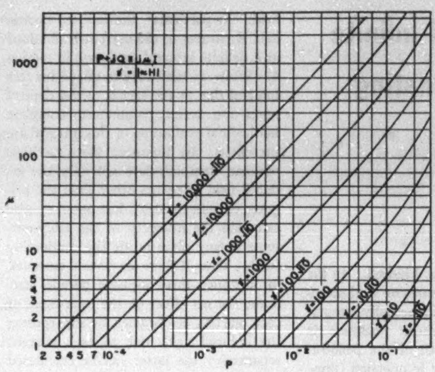


Fig. 2. Impedance function, real part P

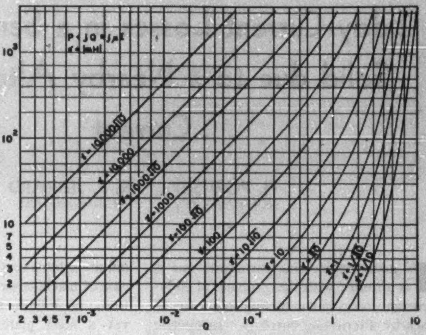


Fig. 3. Impedance function, imaginary part Q

may be neglected, each field component of an alternating electromagnetic field satisfies the differential equation

$$\nabla^2 u = \alpha^2 u \qquad (1)$$

where α^2 is a pure imaginary number, given by

$$\alpha^2 = j4\pi\omega\lambda\mu \times 10^{-9} \qquad (2)$$

and

$j = \sqrt{-1}$, $j^2 = -1$,
f = frequency, cycles per second
$\omega = 2\pi f$ = frequency, radians
μ = permeability
λ = conductivity = $1/\rho$
ρ = resistivity

and where practical centimeter-gram-second (cgs) units are used, so that ρ is in ohm-centimeters. The permeability is dimensionless. Since α^2 is pure imaginary, α lies on the 45-degree axis

$$\alpha = (1+j)\beta; \quad \beta = 2\pi\sqrt{f\lambda\mu \times 10^{-9}} \qquad (3)$$

The real quantity $1/\beta$ has units of length; in fact its units are centimeters, and it is known the "depth of penetration."

The solution of equation 1 shows that a plane-parallel uniform alternating magnetic field, as it penetrates the plane boundary $y = 0$ of a semi-infinite solid $y < 0$ of conductivity λ and permeability μ, dies off in amplitude exponentially in the manner described by

$$e^{\beta y} \text{ where } y < 0 \qquad (4)$$

Both the flux density and the eddy-current amplitudes die off in this manner. It is for this reason that the length

$$d = 1/\beta \qquad (5)$$

is known as the "depth of penetration" for the frequency and for the conductor in question.

In air, $\lambda = 0$ and in place of equation 1,

Table I. Values of P and Q and JμI*

	P+Qi			
log ν = log |αH|	ν=0; μ=1	ν=√10; μ=3.3166248	ν=10; μ=10.0498765	ν=10√10; μ=31.6385840
9/2	0.0^{4}22360 + 0.0^{4}22361j	0.0^{4}741511 + 0.0^{4}741621j	0.0^{3}2246 + 0.0^{3}2247j	0.0^{3}70646 + 0.0^{3}70747j
4	0.0^{4}70701 + 0.0^{4}70711j	0.0^{3}234412 + 0.0^{3}234522j	0.0^{3}7096 + 0.0^{3}7106j	0.0^{2}22276 + 0.0^{2}22368j
7/2	0.0^{3}22351 + 0.0^{3}22361j	0.0^{3}740523 + 0.0^{3}741617j	0.0^{2}22371 + 0.0^{2}22472j	0.0^{2}69760 + 0.0^{2}70734j
3	0.0^{3}706107 + 0.0^{3}707107j	0.0^{2}233426 + 0.0^{2}234516j	0.0^{2}700673 + 0.0^{2}710493j	0.012415 + 0.022333j
5/2	0.0^{2}22261 + 0.0^{2}22361j	0.0^{2}730692 + 0.0^{2}741471j	0.0215069 + 0.0224326j	0.062015 + 0.069798j
2	0.0^{2}69718 + 0.0^{2}70714j	0.0224008 + 0.0234088j	0.0624217 + 0.0699701j	0.15685 + 0.20718j
3/2	0.021383 + 0.022339j	0.0646284 + 0.0730818j	0.158359 + 0.208108j	0.3155 + 0.53621j
1	0.061443 + 0.070023j	0.161203 + 0.216375j	0.315158 + 0.537827j	0.49242 + 1.13383j
1/2	0.147169 + 0.207123j	0.312832 + 0.552089j	0.485277 + 1.131415j	0.62885 + 1.97554j
0	0.256365 + 0.505238j	0.457698 + 1.132140j	0.612087 + 1.961344j	0.70658 + 2.97295j
−1/2	0.334375 + 0.954049j	0.544391 + 1.887395j	0.673702 + 2.924283j	0.7426 + 4.04721j
−1	0.371438 + 1.482351j	0.586270 + 2.714647j	0.719459 + 3.956898j	0.7606 + 5.15234j

	JμI			
	ν=100; μ=100.005000	ν=100√10; μ=316.229347	ν=1000; μ=1000.00050	ν=1000√10; μ=3162.27782
9/2	0.0^{2}222672 + 0.0^{2}223613j	0.0^{2}69725 + 0.0^{2}70696j	0.0214 + 0.0223j	0.06198 + 0.06976j
4	0.0^{2}697283 + 0.0^{2}707000j	0.021405 + 0.022322j	0.0620 + 0.0698j	0.1568 + 0.2071j
7/2	0.021403 + 0.022316j	0.06198 + 0.06976j	0.157 + 0.207j	0.316 + 0.5360j
3	0.06195 + 0.06972j	0.15672 + 0.20708j	0.316 + 0.536j	0.4930 + 1.134j
5/2	0.15682 + 0.20709j	0.31552 + 0.536028j	0.493 + 1.134j	0.6315 + 1.977j
2	0.31552 + 0.53604j	0.493179 + 1.133787j	0.632 + 1.977j	0.7140 + 2.978j
3/2	0.4931 + 1.19236j	0.631545 + 1.977294j	0.714 + 2.981j	0.756 + 4.070j
1	0.6312 + 1.97700j	0.714602 + 2.978204j	0.755 + 4.069j	0.772 + 5.196j
1/2	0.7131 + 2.98045j	0.75478 + 4.068765j	0.773 + 5.195j	0.781 + 6.334j
0	0.7526 + 4.06610j	0.77249 + 5.193782j	0.781 + 6.336j	0.784 + 7.482j
−1/2	0.7688 + 5.18726j	0.77932 + 6.333568j	0.784 + 7.484j	0.784 + 8.633j
−1	0.7788 + 6.32192j	0.78172 + 7.478603j	0.785 + 8.634j	0.784 + 9.784j

* See equations 8 and 10.

each field component satisfies the Laplace equation

$$\nabla^2 u = 0 \qquad (6)$$

By solving equations 1 and 6 with proper continuity conditions at $y=0$, the field of the current configurations of Fig. 1 is obtained. This solution yields for the equivalent impedance Z of the current J flowing at a height h above the plane boundary of the solid $y<0$ the expression

$$Z| = -(2j\omega \times 10^{-9}) \ln \frac{r}{2h} + (j4\mu\omega \times 10^{-9})I \text{ ohms per centimeter (cm)} \qquad (7)$$

where r is the conductor radius and where I is the definite integral

$$I = \int_0^\infty \frac{ed^{-\alpha Hb}b}{\mu b + \sqrt{b^2+1}} \text{ where } H = 2h \qquad (8)$$

while α is given by equation 2.

The impedance Z is the impedance per unit length of conducting wire carrying the current J which would be presented to a generator driving this current, due to the induced field and its eddy currents. (A further slight correction to the right hand member of equation 7 is due to the resistive drop inside the conducting wire carrying the current J and to its linkage of flux inside the wire.)

Since the integral I is a definite integral, its value is independent of the variable of integration b, but it depends on the parameters μ and H, or since α is always a complex number lying on the 45-degree line, on μ and the dimensionless positive number

$$\gamma = |\alpha H| = \sqrt{2}\beta H = 2\sqrt{2}\beta h \qquad (9)$$

Except for the factor $2\sqrt{2}$, γ is the measure of the current distance h from the conductor in units equal to the depth of penetration. The value of I is complex, that is, it has both a real and an imaginary part.

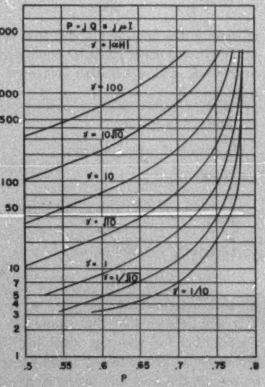


Fig. 4. Impedance function, real part P

The integral I has been evaluated, and is tabulated and plotted as follows. Introducing P and Q

$$P + jQ = j\mu I \qquad (10)$$

one may write the real and imaginary parts of the impedance Z as

$$\text{Im}(Z) = -2\omega \times 10^{-9} \ln \left(\frac{r}{2h}\right) + 4\omega \times 10^{-9} Q \qquad (11)$$

$$\text{Re}(Z) = 4\omega \times 10^{-9} P \qquad (12)$$

Curves of P and Q versus μ for various values of γ are given in Figs. 2 to 5. It is seen that with ω, λ, h, and μ given, the quantity γ may be calculated, and the functions P and Q obtained from these curves. Then, using equations 11 and 12 the resistive and reactive (real and imaginary) components of the impedance Z may be obtained.

More exact values for P and Q are given in Table I. Here, the complex number I is given as a function of $|\alpha H|$ and ν, where $\nu = \sqrt{\mu^2 - 1}$. Interpolation in this table will give more accurate values than can be read from the curves.

As an example, let us suppose that a current-carrying wire is 10 cm from a solid of permeability $\mu = 100$ and resistivity

$$\frac{1}{\lambda} = 80 \text{ microhms per cm} = 0.8 \times 10^{-4} \text{ ohms per cm} \qquad (13)$$

If the frequency is 60 cycles per second, the radian frequency is

$$\omega = 2\pi(60) = 377 \text{ radians per second} \qquad (14)$$

The quantity γ is calculated from equation 9 to be

$$\gamma = 48.66 \qquad (15)$$

Looking at the curves of Fig. 2, one finds by interpolation

$$P = 0.4266 \qquad (16)$$

The resistive component of the impedance is then given by equation 12 and is

$$\text{Re}(Z) = 6.433 \times 10^{-7} \text{ ohm per cm} \qquad (17)$$

If the current has an rms value of 4,000 amperes, the losses per centimeter of length are

$$\text{loss} = (4{,}000)^2 \times 6.433 \times 10^{-7} = 10.29 \text{ watts per cm} \qquad (18)$$

The reactive component of the inductance Z can be calculated in a similar way from equation 7.

2. Field Formulation

Assume that time enters as a factor

$$e^{j\omega t} \qquad (19)$$

in the electric and magnetic fields. Then Maxwell's equations, in cgs practical units, become

$$\nabla \times \mathbf{E} = -j\omega\mu \times 10^{-9}\mathbf{H}$$

$$\nabla \times \mathbf{H} = 0.4\pi\mathbf{i} \qquad (20)$$

where $\mathbf{i}$ is the current-density vector.

Elimination of $\mathbf{H}$ from equations 20 leads to the relation

$$\nabla^2\mathbf{E} = j\omega 4\pi \times 10^{-9}\mathbf{i} \qquad (21)$$

in regions free from charge, provided the following relation be utilized

$$\nabla \cdot \mathbf{E} = 0 \qquad (22)$$

Under the assumption that displacement currents can be neglected, $\mathbf{i}$ is reduced to the conduction current. In regions where the conduction current is due to the induced field, Ohm's law yields

$$\mathbf{i} = \lambda\mathbf{E} \qquad (23)$$

Equation 21 now reduces to

$$\nabla^2\mathbf{E} = \alpha^2\mathbf{E}; \quad \alpha^2 = j4\pi\omega\lambda\mu \times 10^{-9} \qquad (24)$$

Thus each component of $\mathbf{E}$ satisfies the scalar equation 1. Similarly the magnetic-field components can be shown to satisfy equation 1 provided one utilizes the relation

$$\nabla \cdot \mathbf{B} = \mu\nabla \cdot \mathbf{H} = 0 \qquad (25)$$

In free space the conductivity λ may be replaced by zero, and, as noted in sec-

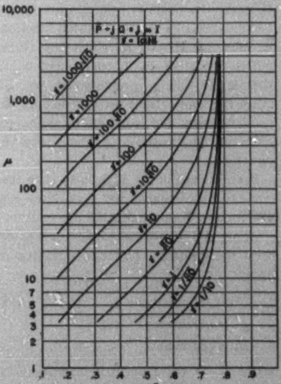


Fig. 5. Impedance function, real part P

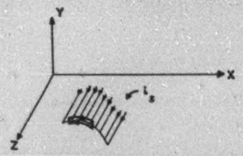


Fig. 6. Curve enclosing current sheet for magnetomotive-force relation

tion 2, equations 1 and 24 for the field components are replaced by the Laplace equation 6.

From now on we consider 2-dimensional fields for which

$$E=(0,0,E_z);\ \partial E_z/\partial z=0 \tag{26}$$

that is fields in which the induced electric field vector points in the direction of the z-axis, while its magnitude depends only on x and y. Such fields arise when the primary currents flow, without change of magnitude or direction, in the direction of the z-axis, and if the conductors, if any, are bounded by cylindrical surfaces whose generators are parallel to the z-axis. No confusion will be caused by denoting E_z by E. Equations 24 and 6 now yield

$$\frac{\partial^2E}{\partial x^2}+\frac{\partial^2E}{\partial y^2}=\alpha^2E \text{ in conductors} \tag{27}$$

$$\frac{\partial^2E}{\partial x^2}+\frac{\partial^2E}{\partial y^2}=0 \text{ in air} \tag{28}$$

The first equation 20 yields

$$(H_x,H_y,H_z)=\frac{10^8}{j\omega\mu}\left(-\frac{\partial E}{\partial y},\frac{\partial E}{\partial x},0,\right) \tag{29}$$

Thus the field can be completely specified in terms of a single scalar function $E(x,y)$ which is a solution of equations 27 and 28. (Usually the fields in question are specified in terms of a vector potential function **A**, such that $\mathbf{H}=\nabla\times\mathbf{A}$, which now turns out to consist of an A_z component only, the latter being proportional to E. Since the conventions used in defining **A** in dielectrics and conductors differ in different treatments, we have used E directly.)

From equation 29, it follows that the curves E=constant are flux lines—as also follows from the flux linkage relation.

Along the cylindrical conductor boundaries the following boundary conditions apply:

1. E is continuous
2. B_n is continuous
3. H_t is continuous

where H_t is the tangential magnetic-field component and B_n the normal flux component. B_n being continuous follows from equation 29 and E being continuous. On the other hand, equation 29 leads to the replacement of H_t being continuous by the boundary condition

4. $(1/\mu)(\partial E/\partial n)$ is continuous

The Poisson equation 21 (as well as its present scalar form) applies to currents distributed over volumes. For surface distributions of currents, as suggested by potential theory, a boundary condition involving a discontinuity in the normal derivative of E applies over the current-carrying surfaces. Thus for current sheets flowing in the z-direction over cylindrical surfaces with generators parallel to the z-axis, the boundary condition applies

$$\Delta\left(\frac{\partial E}{\partial n}\right)=j4\pi\omega\times10^{-9}i_z \tag{30}$$

where i_z is the current density (per unit arc length of a normal section) and the left-hand side denotes the discontinuity in the normal derivative of E. This condition follows from equation 29 and the magnetomotive-force relation

$$\Delta H_t=-(0.4)\pi i_z \tag{31}$$

over the path shown in Fig. 6.

3. Illustrative Examples

The field due to a primary current J following parallel to the plane boundary of a semi-infinite conductor (see Fig. 1) will be obtained in section 4. Several examples required for its solution will now be given. These examples comprise:

1. The field of a single current filament, in air.
2. The field due to a sinusoidal current sheet

$$i_z=A\cos ax \tag{32}$$

flowing over a plane y=constant, in air.

3. The field due the same current sheet located at $y=h$ in the presence of a conductor in the space $y<0$.

Example 1

As the first example we consider

$$E=(j2\omega\times10^{-9})J\ln r$$

$$r=\sqrt{x^2+y^2} \tag{33}$$

from which equation 29 yields

$$H_x=-\frac{0.2Jy}{x^2+y^2};\ H_y=\frac{0.2Jx}{x^2+y^2} \tag{34}$$

or, in polar co-ordinates

$$H_r=0,\ H_\theta=\frac{0.2J}{r} \tag{35}$$

Thus the magnetic lines are circles with the z-axis as axis. The above will be recognized as the field due to a current J flowing along the z-axis, in free space. An arbitrary constant may be added to the right-hand member of equation 33.

We consider next fields of the product form

$$E=\cos ax\,Y(y) \tag{36}$$

where Y is a function of y. Substitution in equations 27 and 28 shows that Y reduces to a linear combination of

$$e^{ay} \text{ and } e^{-ay} \tag{37}$$

in air, and of

$$\exp\left(\sqrt{a^2+\alpha^2}y\right) \text{ and } \exp\left(-\sqrt{a^2+\alpha^2}y\right) \tag{38}$$

in the conductor. The radical $\sqrt{a^2+\alpha^2}$ in equation 38 is complex, and the value which lies in the first quadrant will be taken.

Example 2

As the next example we consider the field, in air, due to the current sheet over the plane $y=0$. This field is given by

$$E=-\frac{j4\pi\omega\times10^{-9}A}{2|a|}e^{-|ay|}\cos ax \tag{39}$$

Indeed the field equation 39 is made up of the product solutions, 36 and 37 so as to avoid infinite values at $y=\pm\infty$, preserve continuity of E at $y=0$, and so as to satisfy the condition of equation 30 for the sinusoidal current, equation 32, over $y=0$

$$\Delta\left(\frac{\partial E}{\partial y}\right)\bigg|_{y=0}=j4\pi\omega\times10^{-9}A\cos ax \tag{40}$$

As a preliminary to example 3, we note that

$$E=j4\pi\omega\times10^{-9}A\times\cos ax\left[\frac{-e^{-|a(y-h)|}+e^{-|a(y+h)|}}{2|a|}\right] \tag{41}$$

represents the field, in air, due to a sinusoidal current sheet given by equation 32 over $y=h$, and a negative image of it, over $y=-h$.

Example 3 (See Fig. 7)

To obtain this field we put

$$E=j4\pi\omega10^{-9}(\cos ax)A\begin{cases}Ce^{-|a|y}-\dfrac{e^{-|a(y-h)|}-e^{-|a(y+h)|}}{2|a|} & \text{for } y>0\\ C\exp\left(\sqrt{a^2+\alpha^2}y\right) & \text{for } y<0\end{cases} \tag{42}$$

where C is a constant.

It will be seen that the component of the solution for $y>0$ not involving C is the same as given by equation 40 and represents the field due to the sinusoidal distribution over $y=h$ along with its negative image at $y=-h$.

For $y<0$, it will be noted that only the first exponential in equation 38 was used in order to avoid infinite values at $y=-\infty$.

Turning to boundary conditions 1 to 4 in section 2 at $y=0$, it will be noted that the first one is satisfied. Application of the second one is facilitated by writing the bracket for $0<y<h$ in the form

$$Ce^{-|a|y}-e^{-|a|h}\times \sinh(|a|y)/|a| \text{ for } 0<y<h \quad (43)$$

Application of condition 4 leads to

$$C=-\frac{e^{-|a|h}}{|a|+\dfrac{\sqrt{a^2+\alpha^2}}{\mu}} \quad (44)$$

Recalling the exponential time factor 19, one may describe the fields of examples 2 and 3 as due to sinusoidal current sheets which pulsate in time. By replacing the sinusoidal factor cos ax by either factor $e^{\pm jax}$, one obtains the fields due to traveling sinusoidal distribution of current. Both of these fields are of interest in electric machine design.

4. Field of a Current Filament Near a Semi-Infinite Solid

By applying Fourier-integral superposition to the fields of example 3 just considered, one may obtain integral representations for fields corresponding to various given current distributions over the plane $y=h$, in the presence of the semi-infinite conducting medium $y<0$. This use of superposition is justified by the fact that the field equations (for instance, equation 21) are linear in the current. For symmetric current distributions $i_s(x)$, this is done by obtaining $A(a)$ from the Fourier resolution of the given current distribution

$$i_s(x)=\int_0^\infty A(a)\cos axda \quad (45)$$

$$A(a)=\frac{1}{\pi}\int_{-\infty}^{\infty} i_s(x)\cos axdx \quad (46)$$

To obtain the corresponding field superposition one multiplies the fields of equations 42 and 44 by da, and integrates over a from $a=0$ to ∞.

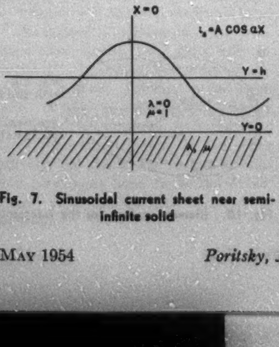


Fig. 7. Sinusoidal current sheet near semi-infinite solid

In particular, for a concentrated current filament J at $x=0$, $y=h$, the "coefficient function" $A(a)$ is given by

$$A(a)=\frac{J}{2\pi} \quad (47)$$

and is independent of a. (The difficulty associated with applying integrations to the highly singular "Dirac function" can be avoided by spreading the concentrated current filament into a current sheet over $|x|<\epsilon$, then letting ϵ approach zero. In the limit, the right-hand member of equation 47 results for $A(a)$.) This follows from equation 46 when $i_s(x)$ is replaced by $J\delta(x)$ where $\delta(x)$ is the "unit-impulse" function or the "Dirac function." One obtains for $y<0$ the field

$$E=-j4\omega\times10^{-9}J\times\int_0^\infty\frac{[\exp(-|a|h+\sqrt{a^2+\alpha^2}y)]\cos axda}{|a|+\dfrac{\sqrt{a^2+\alpha^2}}{\mu}} \text{ for } y<0 \quad (48)$$

For $y>0$, a similar integral results which will be simplified by noting that the portion arising from the exponentials $e^{|a(y-h)|}$ and $e^{|a(y+h)|}$ in equation 42 represents the field, in air, due to the current filament J at $(x,y)=(0,h)$ and its return $-J$ at $(0,-h)$, as follows

$$-j2\omega\times10^{-9}J\int_0^\infty(\cos ax)\times\frac{e^{-|a(y-h)|}-e^{-|a(y+h)|}}{|a|}da=j2\omega\times10^{-9}J\ln(r/r_1) \quad (49)$$

where r is the distance from the current filament J, and r_1 the distance from its image current $-J$ (see Fig. 8). Thus there results

$$E=j4\omega\times10^{-9}J\left[\frac{1}{2}\ln\frac{r}{r_1}-\mu\times\int_0^\infty\frac{e^{-a(y+h)}\cos axda}{\mu a+\sqrt{a^2+\alpha^2}}\right] \text{for } y>0 \quad (50)$$

Equations 48 and 50 constitute the solution of the field in question. Corresponding Fourier integral expressions may be found for H_x and H_y by applying equation 29.

The reader may wonder why it was necessary to drag in the negative image of the sinusoidal current sheet in equation 42, as a result of which the negative image current $-J$ also occurred in the final field equation 50. It is quite true that in place of equation 42 for the field pertaining to Fig. 7, a different expression could be obtained in which the negative image term $e^{-|a(y+h)|}$ is missing. The difficulty with such a procedure is twofold. In the first place, the resulting integrals are nonconvergent, due to the factor a in the denominator; again since the function ln r becomes infinite at infinity, it is hardly a suitable function for Fourier-integral expansion. By including the negative image, both of these difficulties are overcome. Now ln (r/r_1) approaches zero at infinity, and the effect of the vanishing denominator in equations 41 and 49 is overcome by the vanishing numerator.

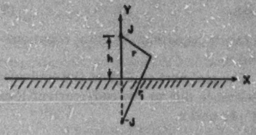


Fig. 8. Current and negative image

Equations 48 and 50 give the field everywhere. In particular, through the relation $i=\lambda E$, equation 48 enables one to calculate the eddy-current density in the conductor $y<0$. Also, by integration of the wattage loss per unit volume $|i^2|/2\lambda$

$$W=\frac{1}{2\lambda}\int_{-\infty}^{0}\int_{-\infty}^{\infty}|i^2|dxdy \quad (51)$$

one obtains the net loss W per unit z due to the eddy currents. However, it is much easier to proceed differently, and to compute the net eddy-current losses from the impedance presented by the induced field at the generator terminals. As shown in Appendix I, this impedance per unit length of primary conductor is given by

$$Z=-\frac{E(0,h)}{J} \quad (52)$$

where as indicated, the numerator represents the value of the induced field E (as given by equation 50) at the current site.

From equations 50 and 52, there results

$$Z=-j2\omega\times10^{-9}\ln\frac{r}{2h}+j4\mu\omega\times10^{-9}\times\int_0^\infty\frac{e^{-2ah}da}{\mu a+\sqrt{a^2+\alpha^2}} \quad (53)$$

This integral is obtained from equation 50 by replacing x by 0 and y by h. Presumably, a similar substitution should be made in the logarithmic term, were it not for the embarassing difficulty of having the term ln r become infinite at $r=0$. This difficulty is due purely to the assumption that the current filament is infinitely thin. For a conductor of finite radius a value of r equal to the radius will be used.

Replacing the variable of integration a by

$$a=b\alpha \tag{54}$$

converts equation 53 into equations 7 and 8.

The real part of Z gives the equivalent resistance due to the eddy-current losses W per unit z

$$W=\frac{|J^2|}{2}\mathrm{Re}(Z) \tag{55}$$

That is, a resistance equal to $\mathrm{Re}(Z)$ per unit length of z, placed in series with the generator, will dissipate as much wattage as the eddy currents do.

As a check, one may obtain the same expression for the loss by integration of the Poynting vector over the surface of the wire; since the only field-energy flow must be into the conductor $y<0$, this result will yield the eddy-current losses, see Appendix II.

When several currents flow simultaneously (for instance, for 3-phase power-line currents) the total induced field can be found by superposition of the individual fields, each one being given by equations similar to equations 48 and 50. Consequently, it is desirable to be able to evaluate the integrals in equations 48 and 50, not only at the current site but also for general x and y. Therefore, while the evaluation of the impedance integral occurring in equation 53 will be discussed primarily in sections 6 to 8, it is also desirable to evaluate the integrals in equations 48 and 50 for general x,y. This is discussed in section 8.

5. Discussion of the Field and Impedance Integrals

First we shall simplify the integrals of equations 50 and 53 by introducing the change of variable, equation 54, and expressing cos ax in terms of imaginary exponentials. The integral in equation 50 is expressed as a sum of two integrals of the form of equation 8

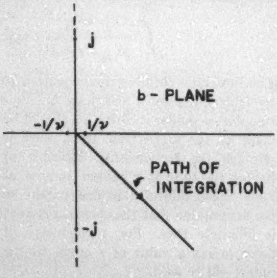


Fig. 9. Path of integration and location of singularities

$$I=\int_0^\infty \frac{e^{-\alpha Hb}db}{\mu b+\sqrt{b^2+1}} \tag{56}$$

where

$$H=h+y\pm jx \tag{57}$$

Similarly, the integral in the impedance equation 53 is reduced to the form of equation 56, with $H=2h$. Throughout the remainder of the paper the letter I will denote the integral of equation 56.

Denote by $f(b)$ the coefficient of the exponential of the integrand in equation 56

$$f(b)=\frac{1}{\mu b+\sqrt{b^2+1}} \tag{58}$$

The singularities of $f(b)$ may be turned to account to give a clue to advantageous breakdown of I. These singularities consist of branch points at

$$b=\pm j \tag{59}$$

and poles at the roots of

$$\mu b+\sqrt{b^2+1}=0 \tag{60}$$

namely at

$$b=\pm\frac{1}{\nu};\quad \nu=\sqrt{\mu^2-1} \tag{61}$$

When μ is large, the poles are close to the origin; as μ decreases, the poles move away from the origin and cross the unit circle when $\mu=\sqrt{2}$. As μ approaches unity, the poles move toward ∞, and they disappear altogether when $\mu=1$. These singularities are shown in Fig. 9 for μ about 10.

To represent the double-valued function $f(b)$, one may use a "Riemann surface" consisting of two sheets which are joined to each other along certain lines or curves. The "upper" sheet may be considered to carry the values of the integrand for which the real part of $\sqrt{b^2+1}$ is positive; the "lower" sheet the values for which it is negative. The two sheets will then be joined along curves where $\sqrt{b^2+1}$ is a pure imaginary; that is, along two paths which consist of the imaginary axis from $b=j$ to $b=j\,(+\infty)$ and the imaginary axis from $b=-j$ to $b=j(-\infty)$ (these are shown in Fig. 9 as broken lines). The Riemann surface is shown schematically in Fig. 10. A path on this Riemann surface passes from one sheet to the other whenever it crosses either of the junction lines.

It is readily seen from inspection of the integrand of I that the pole $b=-1/\nu$ lies on the upper Riemann sheet, while the pole $b=1/\nu$ lies on the lower sheet. The path of integration on the b-plane corresponding to integration along the real axis in equations 48 and 50 is the -45-degree line, and lies in the upper sheet. Therefore, for large values of μ the pole $b=-1/\nu$ lies very close to the path of integration and will thus contribute an appreciable part of the value of I; while $b=1/\nu$, in spite of its appearance in Fig. 9, lies "far away" from this path.

The residue of $f(b)$ at $b=-1/\nu$ is

$$\frac{\mu}{\nu^2} \tag{62}$$

where, as in equation 61

$$\nu=\sqrt{\mu^2-1} \tag{63}$$

Therefore the function

$$f_1(b)=\frac{\mu}{\nu^2}\frac{1}{b+\frac{1}{\nu}}=\frac{\mu}{\nu}\frac{1}{b\nu+1} \tag{64}$$

has the same singularity at the pole $b=-1/\nu$ as $f(b)$. The resolution

$$f(b)=f_1(b)+f_2(b) \tag{65}$$

where

$$f_2(b)=-\frac{1}{\nu^2}\frac{1}{\frac{\mu}{\nu}+\sqrt{b^2+1}}=\frac{\mu/\nu-\sqrt{b^2+1}}{\nu^2b^2-1} \tag{66}$$

is analytic at $b=-1/\nu$, and breaks I down into two integrals

$$I=I_1+I_2 \tag{67}$$

where

$$I_1=\int_0^\infty f_1(b)e^{-\alpha Hb}db=\frac{\mu}{\nu}\int_0^\infty \frac{e^{-\alpha Hb}}{b\nu+1}db \tag{68}$$

$$I_2=\int_0^\infty f_2(b)e^{-\alpha Hb}db=\frac{1}{\nu^2}\times \int_0^\infty \frac{e^{-\alpha Hb}db}{\frac{\mu}{\nu}+\sqrt{b^2+1}} = \int_0^\infty \frac{\frac{\mu}{\nu}-\sqrt{b^2+1}}{\nu^2b^2-1}e^{-\alpha Hb}db \tag{69}$$

The integrand of I_1, it will be noted, has no branch points. I_1 can be evaluated in terms of what is known as the "integral-exponential" function Ei, denoted in the

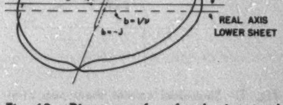


Fig. 10. Riemann surface for the integrand

following by E_1. The integral I_2, on the other hand, presents a rather difficult integration. The method used in the following for its evaluation consists in expanding $f_2(b)$, first in a series of positive powers

$$f_2(b)=\sum_{n=1}^{\infty} D_{2n}b^{2n} \text{ for } |b|<1 \tag{70}$$

valid inside the unit circle, then in a negative power series

$$f_2(b)=\sum_{n=1}^{\infty}\frac{G_n}{b^n} \text{ for } |b|>1 \tag{71}$$

valid outside the unit circle. The integral I_2 is now broken up into two parts as follows

$$I_2=\int_0^{\infty}=\int_0^{b_0}+\int_{b_0}^{\infty}=I_A+I_B \quad \text{for } |b_0|=1 \tag{72}$$

where b_0 lies on the unit circle. Substitution from equations 70 and 71 yields

$$I_A=\sum D_{2n}\int_0^{b_0} b^{2n}e^{-(\alpha H)b}db \tag{73}$$

$$I_B=\sum G_n\int_{b_0}^{\infty}\frac{e^{-(\alpha H)b}}{b^n}db \tag{74}$$

For real αHb, the integrals in equations 73 and 74 can be reduced to forms for which tables exist.

It turns out that for large values of μ and large values of $H\beta$, the integral I_1 presents a sufficiently close approximation to the value of I. Indeed, under these conditions the derivative $(\partial^2E)/(\partial x^2)$ in the differential equation 27 can be neglected in comparison with the derivative $(\partial^2E)/(\partial y^2)$. Then the differential equation 27 can be integrated completely for each x in terms of its boundary value and a proper application of conditions 1 and 4 of section 2 can be used to eliminate the solution of the field in $y<0$, and replace equation 27 by a proper boundary condition at $y=0$ involving the values of E and $(\partial E)/(\partial y)$ at $y=0$. For details of this approximate procedure, the reader is referred to reference 4. Application of this method to the present problem leads precisely to a field expression in which the integral I is replaced by I_1, but with ν replaced by μ.

The method of calculation briefly outlined above is explained more fully in section 7. It sufficed for most of the cases considered. For large μ and large $|\alpha H|$, certain simplifications could be effected in the calculation. On the other hand, for small $|\alpha H|$ some of the series converge too slowly, and further calculation procedures have to be developed. These modifications are discussed in section 7.

6. Evaluation of the Integrals I₁ and I₂

To reduce the integral I_1 of equation 64 we introduce the variable of integration

$$u=\alpha H\left(b+\frac{1}{\nu}\right) \tag{75}$$

whereupon there results

$$I_1=\frac{\mu}{\nu^2}e^{\frac{\alpha H}{\nu}}E_1\left(\frac{\alpha H}{\nu}\right) \tag{76}$$

where E_1 (also sometimes denoted by Ei) is defined by

$$E_1(z)=\int_z^{\infty}\frac{e^{-u}}{u}du \tag{77}$$

Two well-known expansions are available for evaluation of this integral. The power-series expansion of e^{-u}, followed by division by u and integration yields

$$E_1(z)=-\ln|z|-i\theta-C+\sum_{n=1}^{\infty}\frac{(-1)^{n+1}z^n}{(n)(n!)} \tag{78}$$

where $C=0.57722$.

This converges for all z, but slowly for large $|z|$. The other expansion is obtained by integration by parts, which yields the asymptotic series

$$E_1(z)=e^{-z}\left[\frac{1}{z}-\frac{1}{z^2}+\frac{2!}{z^3}-\frac{3!}{z^4}+\dots\right] \tag{79}$$

Though divergent, this series is useful for large $|z|$, just where the series (78) converges too slowly. Tables of the function $E_1(z)$, where $z=x+iy$, for $|x|<3$ and $|y|<3$ are given in reference 5.

We turn next to the integral I_2 and its expansion, equations 69 and 70. First, consider a direct power series of $f(b)$, the coefficient of $e^{-\alpha Hb}$ in the I-integral, see equations 56 and 58

$$f(b)=\sum_{n=0} C_nb^n \tag{80}$$

This was given by Wise.[2] To obtain it one expands $f(b)=[\mu b+(b^2+1)^{1/2}]^{-1}$ by the binomial theorem in positive powers of b, then further expands the negative powers of $(b^2+1)^{1/2}$, and collects like powers of b. There results for odd n, see equation 63

$$-C_n=\mu(\mu^2-1)^{(n-1)/2}=-\mu\nu^{n-1} \tag{81}$$

while for even n, C_n is equal to the expansion of the second member of equation 81 in descending powers of μ, stopping with the term which is independent of μ. We shall use a double bracket to denote such a sum. Hence, for even n

$$C_n=[[\mu(\mu^2-1)^{(n-1)/2}]] \tag{82}$$

Next we expand the function $f_1(b)$, see equation 64, in powers of b obtaining

$$f_1(b)=\frac{\mu}{\nu}\frac{1}{\nu b+1}=\frac{\mu}{\nu}\times(1-\nu b+\nu^2b^2-\nu^3b^3+\dots) \tag{83}$$

Subtracting the series, equations 80 and 83, one notes that the odd coefficients cancel, while in the even coefficients the terms with integral powers of μ and the constant terms (free from μ) cancel. There results

$$D_{2n}=C_{2n}-\mu\nu^{2n-1}=[[\mu(\mu^2-1)^{(2n-1)/2}]]-\mu(\mu^2-1)^{(2n-1)/2}=(-1)^n\times\left[\left(\frac{2n-1}{2},n+1\right)/\mu^2-\left(\frac{2n-1}{2},n+2\right)\times/\mu^4+\dots\right] \tag{84}$$

where (m,n) are the binomial coefficients, as explained in the following. Recurrence relations for the coefficients D_{2n} are found to be

$$D_0=1-\mu\nu=\frac{-1}{\nu(\mu+\nu)};\quad D_2=\nu^2D_0+\frac{1}{2};\dots$$

$$D_{2n}=\nu^2D_{2n-2}+\left(\frac{1}{2},n\right)=\nu^2D_{2n-2}+(-1)^{n-1}\frac{(2n-3)!!}{2n!!} \tag{85}$$

where $(m)n$, denotes, as usual, the binomial coefficient; that is, the coefficient of x^n in the expansion of $(1+x)^m$ in powers of x. The double factorial notation $n!!$ denotes the product of every other integer in decreasing order, starting with n; thus, $(2n)!!=2\cdot4\cdots(2n)$; $(2n+1)!!=1\cdot3\cdots(2n+1)$; furthermore, $0!!=1$, $(-1)!!=1$.

With D_{2n} evaluated we turn to the integrations in equation 73

$$\int_0^{b_0}e^{-(\alpha Hb)}b^ndb \tag{86}$$

where b_0 as shown in Figs. 9 and 10 is on the -45-degree axis and on the unit circle

$$b_0=e^{-j\pi/4}=\frac{1-j}{\sqrt{2}} \tag{87}$$

Since for real H, the argument of the exponential (αHb) is real, we make the change of variables

$$\alpha Hb=s \tag{88}$$

and obtain

$$\left[\int_0^{\beta H\sqrt{2}}e^{-s}s^nds\right](\alpha H)^{n+1} \tag{89}$$

Integration by parts yields

$$\int_0^x e^{-s}s^nds=n\int_0^x e^{-s}s^{n-1}ds-e^{-x}x^n \tag{90}$$

and hence, by induction

$$\int_0^x e^{-s}s^n ds = n!\left[1-e^{-x}\left(1+x+\frac{x^2}{2!}+\ldots+\frac{x^n}{n!}\right)\right] \quad (91)$$

These equations suffice for the integrations in equation 73.

Alternatively one may use the values of the ratio

$$\frac{\int_0^x e^{-s}s^n ds}{\int_0^\infty e^{-s}s^n ds}=\frac{\int_0^x e^{-s}s^n ds}{n!} \quad (92)$$

tabulated by Karl Pearson,[6] though what is tabulated are the values of

$$I(u,p)=\int_0^{\sqrt{(p+1)}u} e^{-v}v^p dv/p! \quad (93)$$

In terms of these "incomplete Γ-functions" one obtains

$$I_A=\sum_{n=0}^{\infty}\frac{D_{2n}(2n)!I\left(\beta H/\sqrt{n+\frac{1}{2}},\ 2n\right)}{(\alpha H)^{2n+1}} \quad (94)$$

We now turn to the evaluation of the coefficients G_n in the expansion and integrations of equation 74.

Evaluation of G_n is carried out by expanding $\sqrt{b^2+1}$ in the numerator on the right-hand side of equation 66 in negative powers of b, substituting the series, equation 71, for $f_2(b)$ in equation 66 and clearing of fractions

$$(\nu^2b^2-1)\left(\sum G_n/b^n\right)=\left\{\frac{\mu}{\nu}-\left[b+\left(\frac{1}{2},1\right)b^{-1}+\left(\frac{1}{2},2\right)b^{-3}+\ldots\right]\right\} \quad (95)$$

Equating coefficients of like powers of b there results

$$G_{2n}=\frac{\mu}{\nu^{2n+1}} \quad (96)$$

for the even coefficients and the recursion formulas

$$G_1=-\frac{1}{\nu^2};\quad G_3=-\frac{\left(\frac{1}{2},1\right)}{\nu^2}-\frac{1}{\nu^4};\quad \ldots \quad (97)$$

$$G_{2n+1}=-\frac{\left(\frac{1}{2},n\right)}{\nu^2}+\frac{1}{\nu^2}G_{2n-1} \quad (98)$$

for the odd coefficients.

Utilizing the same change of variable equation as in the foregoing in equation 74 one obtains

$$I_B=\sum_1^\infty(\alpha H)^{n-1}G_n\int_{|\alpha H|}^\infty\frac{e^{-u}du}{u^n} \quad (99)$$

A final change of variable

$$u=|\alpha H|x \quad (100)$$

yields

$$I_B=\sum_1^\infty G_n\exp\left[j\left(\frac{n-1}{4}\pi\right)\right]\times E_n(|\alpha H|) \quad (101)$$

where the functions E_n $(|\alpha H|)$ are identified as the functions $E_n(x)$ of reference 7.

The method of calculation of this and the preceding section was applied to obtain the values entered in Table I. There was no trouble in evaluating the integral I_1. Likewise, the series for I_A in equation 93 was generally found to converge rapidly. When I_B was calculated by equation 96 the series was found to converge much less rapidly than equation 94. This is due to the slow rate of decrease in size of the G_{2n+1} and also of the functions $E_n(x)$.

The error committed in breaking off the series, equation 101, at an odd term is of the order of magnitude $\sqrt{2}$ times the modulus of that term. The error in breaking off at an even term is not more in magnitude than if the series had been broken off at the preceding term.

The functions $E_n(x)$ of low order are relatively large for $x=1$ and even for $x=\sqrt{10}$. But they are quite small for $x=10$. Thus, when ν and $|\alpha H|$ are both under 10, the computation of I_B requires quite a few terms of the series, equation 96, for high accuracy. On the other hand, when $|\alpha H|\geqq 10\sqrt{10}$, I_B supplies a negligible contribution, so that I_A may be taken alone for I_2.

From the results obtained by the method of this section one may draw the following conclusions:

1. For $|\alpha H|\leqq 10$, $\nu=10$ or $\sqrt{10}$, computation of I by the method of this section is indicated: $I=I_1+I_A+I_B$.
2. For $10\sqrt{10}=\leqq|\alpha H|$, $\nu=10$ or $\sqrt{10}$, it will be found sufficiently accurate for many purposes to take $I=I_1+I_A$.
3. For $10\sqrt{10}\leqq\nu$, the approximation $I=I_1$ will be adequate, except for situations demanding extreme accuracy.

7. Simplification in Procedure and Alternative Methods. Asymptotic Series (Approximation for Large μ)

For μ large, equation 66 shows that $f_2(b)$ can be approximated as follows

$$f_2(b)=-\frac{1}{\nu^2}\frac{1}{\frac{\mu}{\nu}+\sqrt{b^2+1}}\cong-\frac{1}{\nu^2}\frac{1}{1+\sqrt{b^2+1}}=\frac{1}{\nu^2}\frac{1-\sqrt{b^2+1}}{b^2} \quad (102)$$

and I_2 reduces to

$$I_2'\cong\frac{1}{\nu^2}\int_0^\infty\left(\frac{1-\sqrt{b^2+1}}{b^2}\right)e^{-(\alpha H)b}db \quad (103)$$

where the integral is no longer dependent on μ.

The evaluation of the integral of equation 103 can be carried out by similar methods. In particular, the expansion of equation 70 now becomes

$$D_{2n}\rightarrow D_{2n}'=\frac{(-1)^{n+1}}{2\nu^2}\frac{(2n-1)!!}{2^n(n+1)!};\quad D_{2n+1}=0,\ D_0=-1/2\nu^2 \quad (104)$$

The error I_2-I_2' can be estimated not to exceed $1/[4\nu^3(\mu+\nu)\beta H]$ in magnitude.

As pointed out in section 7 for small $|\alpha H|$ (such as $|\alpha H|=0.1$), the convergence of the series of equation 100 was exceedingly poor. To remedy this, the range of b-integration from b_0 to ∞ was broken up at $|b|=1.5$. This improved the convergence of the E_n-series which replaced equation 100. Over the interval between $|b|=1$ and $|b|=1.5$, the exponential $e^{-\alpha Hb}$ was expanded in positive powers of b and the product of this series by the negative power series of equation 71 was found and integrated term by term.

It turned out that for some n and for small values of βH, linear and even quadratic interpolation in the tables of reference 6 proved inaccurate; in these cases, equation 91 was resorted to, to evaluate the integrations in equation 74.

Wise[2] obtains an asymptotic expansion for I by substituting the series of equation 80 for $f(b)$ in equation 56 directly and integrating term by term, obtaining

$$I=\sum_{n=0}^\infty\frac{n!\,C_n}{(\alpha H)^{n+1}} \quad (105)$$

This divergent series is useful for values of μ close to 1, provided $|\beta|$ is large. However, for large μ, the convergence of the series of equation 80 is limited by the pole $b=-1/\nu$, and as a result, except for fantastically large $|\alpha H|$, the series of equation 105 is relatively useless. However, after removal of the pole from I by subtracting I_1, one may try using the power series of equation 70 in I_2, with termwise integration from 0 to ∞. This yields

$$I_2=\sum_{n=0,2,\ldots}\frac{D_n\,n!}{(\alpha H)^{n+1}} \quad (106)$$

which is very useful, even for large μ, again provided that $|\alpha H|$ is large.

8. Power-Series Expansions for the Field and Impedance Integrals

Throughout the preceding sections, real values of H were implied, and to render αHb, the argument of the exponential in

I, I_1, and I_2, real, the -45-degree path of integration was used. If, as in evaluating E away from the current site, H is complex, as in equation 57, then the path of integration may be shifted from the -45-degree line to the line

$$\arg(b) = -\arg(\alpha H) \qquad (107)$$

with proper modifications in the terms of equation 101.

Some distinct advantages accrue from choosing the path of integration so that it deliberately passes through a branch point, $b=-j$. This was probably the method used in reference 1 in evaluating the field integrals for the case $\mu=1$. Indeed for $\mu=1$, equation 56 becomes

$$I=\int_0^\infty \frac{e^{-\alpha H b}db}{b+\sqrt{b^2+1}} \qquad (108)$$

and rationalizing the denominator as in equation 66 one obtains

$$I=-\int_0^\infty be^{-\alpha H b}db+\int_0^\infty \sqrt{b^2+1}e^{-\alpha H b}db$$
$$=\frac{-1}{(\alpha H)^2}+\int_0^\infty \sqrt{b^2+1}e^{-\alpha H b}db \qquad (109)$$

Changing the variable of integration to c

$$b=ic \qquad (110)$$

the branch points move to $c=\pm 1$. Choosing the path of integration to pass through 1, the integral in equation 109 becomes

$$\int_0^1 \sqrt{c^2-1}e^{-(\alpha H i)c}dc+$$
$$\int_1^\infty \sqrt{c^2-1}e^{-(\alpha H i)c}dc \qquad (111)$$

The second integral in equation 111 can be evaluated in terms of the Bessel function $H_1^{(2)}(\alpha H)$, see Watson,[7] pages 169, 170. As regards the first integral, it can be evaluated as a power series in (αH) by expanding the exponential, integrating termwise, and utilizing the β-integral

$$\int_0^1 u^\alpha(1-u)^\beta du=\frac{\alpha!\beta!}{(\alpha+\beta)!}$$

where

$$u=c^2$$

and

$$\alpha!=\Gamma(\alpha+1) \qquad (112)$$

Actually the series arising from the even terms of the exponential can be reduced to the Bessel function $J_1(\alpha H)$, see reference 7, page 48, equation 31, while the odd terms lead to a Struve function, see reference 7, page 328; the latter, however, has not been tabulated for complex arguments, and the direct use of its series is recommended.

By expanding $H_0^{(2)}$ in well-known convergent series, see reference 7, pages 62, 64, and 73 one expresses the integral in equation 109 in the form

$$\int_0^\infty \sqrt{b^2+1}e^{-\alpha H b}db=\frac{P(\alpha H)}{(\alpha H)^2}+$$
$$Q(\alpha H)\ln(\alpha H) \qquad (113)$$

where P and Q are power series in (αH), see reference 1, pages 8–9. The same form applies also to the complete integral I of equation 109.

It will now be shown how series similar to equation 113 can be obtained for I for any μ. This series for I_1 is given by equations 76 and 78. Hence, only the series for I_2 remains to be considered.

Note that by applying the operator

$$\left[\nu^2\frac{\partial^2}{\partial(\alpha H)^2}-1\right] \qquad (114)$$

to I_2 in the form given by the last term in equation 69, the denominator is cancelled and one obtains

$$\nu^2\frac{\partial^2 I_2}{\partial(\alpha H)^2}-I_2=\int_0^\infty\left(\frac{\mu}{\nu}-\right.$$
$$\left.\sqrt{b^2+1}\right)e^{-\alpha H b}db=\frac{\mu}{\nu}\frac{1}{\alpha H}-$$
$$\int_0^\infty\sqrt{b^2+1}e^{-(\alpha H)b}db \qquad (115)$$

The last integral is precisely the integral occurring in equation 109 for the case $\mu=1$ and for which the representation of equation 113 holds. Hence, I_2 can be evaluated by solving the differential equation

$$\nu^2\frac{\partial^2 I_2}{\partial(\alpha H)^2}-I_2=\frac{P(\alpha H)}{(\alpha H)^2}+Q(\alpha H)\ln(\alpha H)$$
$$(116)$$

The contribution toward I_2 arising from the $(\alpha H)^{-2}$ and $(\alpha H)^{-1}$ terms can be expressed in terms of E_1-functions. There remains to integrate for the contribution to I_2 arising from the solution of the equation $s=\alpha H$

$$\nu^2\frac{\partial^2 u}{\partial s^2}-u=P(s)+Q(s)\ln s \qquad (117)$$

Assume a solution of equation 117 exists of the form

$$u=p(s)+q(s)\ln s \qquad (118)$$

where p and q are to be presently specified. Substituting the form of equation 118 in equation 117, and equating coefficients, leads to

$$\nu^2q''(s)-q(s)=Q(s) \qquad (119)$$

$$\nu^2p''(s)-p(s)=P(s)-\nu^2\left[\frac{2q'(s)}{s}-\frac{q}{s^2}\right] \qquad (120)$$

Equation 119 can be integrated for $q(s)$ as a power series whose coefficients depend linearly on those of $Q(s)$. Likewise, there exists a power-series solution of equation 120 for p corresponding to the non-negative powers on the right-hand side. The latter, again, contribute terms expressible in terms of the E_1-function. In summary, it appears that a representation similar to the right-hand member of equation 113 obtains for I_2 for any μ.

Thus far only the field in air (given by equation 50) was discussed. We now consider briefly the integral of equation 48 for the field in the conducting solid. Upon introduction of b, as in equation 54, and making the further change of variable

$$b=\frac{1}{2}\left(u-\frac{1}{u}\right) \qquad (121)$$

it is found that the radical disappears in the integral of equation 48 and it can be broken up into a sum of integrals of the form

$$\int_1^\infty \frac{(u^2-1)\{\exp[(Au+B/u)/2]\}du}{u^2(1+\mu)+(1-\mu)}$$

where

$$A=\alpha(-h+y\mp ix)$$

and

$$B=\alpha(h+y\pm ix) \qquad (122)$$

Possible methods of evaluation of this integral have been studied. One may utilize the expansion $\exp[B/u]=\Sigma B^n/(n!u^n)$ for the distant portion of the path of integration and the Laurent series, see reference 7, page 14

$$\exp[(Au+B/u)/2]$$
$$=\sum\left(u\sqrt{\frac{A}{B}}\right)^n I_n(\sqrt{AB}) \qquad (123)$$

for the main part of the integration path. The exponential factor is broken up into partial fractions and the poles removed by subtraction of rational terms, which lead to logarithmic integrals, and Laurent series can then be found for the remainder, and integrated termwise.

9. Representation of the Field by Means of Images

We now consider the electric field in air and, in particular, the part of it given by the integral in equation 50. If the integrand of this integral is broken up as in equation 65 a sum of two integrals I_1 and I_2, results, of which I_1 can be integrated as in equation 76. It is shown in reference 4, that the field component due to I_1 can be represented by means of a positive-current image at $(x,y)=(0,-h)$ and an exponential trail of images, extending along the negative y-axis from $y=-h$ to $y=-\infty$.

A somewhat similar conclusion can be established for the field component due to the integral I_2. If this integral be denoted by

$$E=\int_0^\infty D(a)e^{-ay}\cos axda \qquad (124)$$

then it can be shown that this integral may be transferred into

$$2\int_0^\infty I(s)\ln r ds,\; r^2=x^2+(y+s)^2 \qquad (125)$$

where $I(s)$ represents the image current distribution and is given by the operational expression

$$I(s)=\frac{pD(p)}{2}p\mathbf{1}(s)=\frac{p^2D(p)}{2}\mathbf{1}(s) \qquad (126)$$

or by the Bromwich-integral form

$$I(s)=\frac{1}{4\pi i}\int pD(p)e^{ps}dp \qquad (127)$$

Here the path of integration is parallel to the pure imaginary axis in the p-plane and lies to the right of the singularities of $D(p)$.

It will be noted that except for a factor p and the difference in the path of integration, the integral in equation 127 is the same as the field integral for the field, equation 124. Its evaluation has been studied sufficiently to verify the following results:

1. The current distribution $I(y)$ vanishes for $-h<y<0$.
2. $I(y)$ possesses a concentrated current at $y=-h$.
3. $I(y)$ consists of a continuous distribution of current images extending from $y=-h$ to $y=-\infty$.

Appendix I

The electromagnetic-force equation in the generator primary-current circuit (neglecting the resistance of the wire carrying the current J) is

$$E_{\text{appl}}+E_{\text{ind}}=Z_iJ \qquad (128)$$

where E_{appl} is the applied, or generator, voltage; E_{ind} the voltage induced along the current path, and Z_i the internal generator impedance. If the voltages in equation 128 are calculated per unit z, then E_{ind} is given by E from equation 50 evaluated at $x=0$, $y=h$. Since E is proportional to J, we may take out J as a factor and write E in the form

$$E|_{x=0,\,y=h}=-ZJ \qquad (129)$$

where Z is a proper constant, independent of J. Equation 128 then yields

$$J=\frac{E}{Z_i+Z} \qquad (130)$$

Hence, from point of view of the generator, the field acts as if an impedance Z per unit z were inserted in series with internal generator impedance. Equation 129 is the same as equation 52. The explicit form of Z is given by equations 7 and 53.

The reader may find it difficult to conceive of the return of the current to the generator terminals when the fields studied allow of no such return for any currents. Now, it can be shown that the net eddy-current resultant in the conductor $y<0$ is equal to $-J$. Hence, if one generator terminal is grounded near the generator, while the other one leads to the conductor carrying the current J of great length l and whose far end is also grounded, the impedance viewed by the generator will be lZ, except for an "end-effect" correction, which may be made relatively small by making l large.

Appendix II

The radial component of the Poynting vector $\mathbf{E}\times\mathbf{H}$ is to be integrated over the surface of the conductor carrying the primary current J. The tangential magnetic field, it will be noted from equation 29 is proportional to the normal (radial) derivative of E, $\partial E/\partial n$. Hence, only the electric field component due to ln r, yields a nonvanishing result. The magnetic field due to the latter is given by equation 35 and leads readily to an energy flow identical with equation 55.

References

1. WAVE PROPAGATION IN OVERHEAD WIRES WITH GROUND RETURN, **John R. Carson.** *Bell System Technical Journal*, New York, N. Y., vol. V, 1926, pp. 538–54.

2. EFFECT OF GROUND PERMEABILITY ON GROUND RETURN CIRCUITS, **W. Howard Wise.** *Bell System Technical Journal*, New York, N. Y., vol. X, 1931, pp. 472–84.

3. TRAVELING WAVES ON TRANSMISSION SYSTEMS (book), **L. V. Bewley.** John Wiley and Sons, Inc., New York, N. Y., 1951.

4. ON REFLECTION OF SINGULARITIES OF HARMONIC FUNCTIONS CORRESPONDING TO THE BOUNDARY CONDITION $\partial u/\partial n+au=0$, **H. Poritsky.** *Bulletin of American Mathematical Society*, New York, N. Y., Dec. 1937, pp. 881–82.

5. TABLE OF $E(z)=e^z\int_z^\infty \frac{e^{-u}du}{u}$ FOR SPECIAL VALUES OF x AND y AND TABLE OF $E_1(z)=\int_z^\infty \frac{e^{-u}du}{u}$ Mathematic Tables Project, New York, N. Y.

6. TABLES OF THE INCOMPLETE Γ-FUNCTION (book), **Karl Pearson,** editor. London, England, 1922.

7. THEORY OF BESSEL FUNCTIONS, (book), **G. N. Watson.** The MacMillan Company, New York, N. Y., second edition, 1944.

◆

No Discussion

Polyethylene-Insulated Telephone Cable

A. S. WINDELER
NONMEMBER AIEE

Synopsis: The physical properties of polyethylene are such as to make it attractive for many wire-insulating applications, particularly in multiconductor communications cables. This article presents certain factual information relating to new types of multiconductor cables having extruded polyethylene insulation, and describes briefly their initial installation in the working telephone plant. The literature is replete with information on the physical and chemical properties and the behavior of polyethylene, and so no attempt is made to explore the quality of the material *per se.* Polyethylene insulation extruded in the form of both solid material and foam to impart certain desired electrical properties is discussed. In a broad sense, this article may be considered as announcing an important new insulating material for telephone cables, which may be expected eventually to have very extensive applications in the Bell System plant.

FROM almost the beginning of the art, multiconductor telephone cables have been insulated with paper, applied as a helical tape or laid down directly on the conductor in the form of pulp. Solid paper has a dielectric constant in the order of 2.5 to 3.0, but in the case of either ribbon or pulp insulation, a considerable amount of air is included in the electric field surrounding the conductor, so that the composite effective dielectric constant is of quite low value, usually about 1.5 to 1.6 in a typical design. In recent years, as various plastics and other polymeric materials have become available, these have been studied as competitors of paper, and polyethylene in particular now appears to have an important field of application. Polyethylene appears attractive because of its excellent electric properties, including low dielectric constant and power factor, compared with other usable plastics, and high dielectric strength. It is also highly imper-

---

Paper 54-104, recommended by the AIEE Wire Communications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 27, 1953; made available for printing December 21, 1953.

A. S. WINDELER is with the Bell Telephone Laboratories, Inc., Baltimore, Md.

Table I. Characteristics of Insulating Materials

	Polyethylene	Plasticized Polyvinylchloride	Polystyrene	Polytetrafluoroethylene (Teflon)	Polyamide (Nylon)
Density, grams per cubic centimeter	0.92	1.2 to 1.4	1.06	2.2	1.09
Tensile strength, pounds per square inch	1400 to 2000	1500 to 3000	500 to 9000	1500 to 2500	7000
Elongation, per cent	600	200 to 450	2 to 5	100 to 200	100 to 200
Water absorption, per cent in 24 hours	<0.01	0.4 to 0.65	<0.05	nil	0.4 to 2.0
Dielectric strength, rms volts per mil 1/8-inch thickness*	400 to 500	300 to 700	500 to 700	400 to 500	400
Power factor, 1 to 300 kc	0.0002	0.09 to 0.16	0.0002	0.0002	0.04 to 0.2
Dielectric constant, 1 to 300 kc	2.3	3.5 to 5.0	2.5 to 2.6	2.0	3.5 to 8

* Dielectric strengths are greater for thinner sections; for example, in 14-mil thicknesses, polyethylene has a dielectric strength of approximately 2,500 rms volts per mil.

meable to water or water vapor and is available in the desired quantities at a reasonable price. In addition, it is considered probable that the long-term price trend will be downward.

The various electrical and mechanical characteristics[1,2] of polyethylene are shown in Table I, along with some of the other materials considered. Among these other materials only polytetrafluoroethylene (Teflon) and polystyrene have power factors and dielectric constants in the same low range as polyethylene. There are basic objections to both of these materials. Polytetrafluoroethylene is so expensive as to be uneconomical for this application, and polystyrene in thick sections is too stiff and brittle to handle in a satisfactory manner.

The use of polyethylene in telephone cables is not altogether new, but it has heretofore been confined to special types of high-frequency cable. For example, the coaxial cable and the video pair have polyethylene-disk insulation and ribbon-and-string insulation, respectively, see Fig. 1. These cables were designed for low attenuation in the megacycle range, and the use of polyethylene or a similar low-power-factor material was a necessity. A low power factor is of lesser importance in the carrier systems for which the multipair cables are used, and the polyethylene insulation on these cables must show other advantages to prove desirable.

Fig. 1. High-frequency cable

A. Polyethylene-disk insulated coaxial
B. Polyethylene ribbon-and-string-insulated video pair

Note expanded polyethylene interstice fillers

Polyethylene insulation is applied to the wire by an extrusion process; the insulation may be either solid or expanded depending on the application. Generally the polyethylene is supplied as granules previously compounded with a antioxidant. In the case of solid polyethylene insulation, the granules, thoroughly mixed with pigment, are fed into the extruder and formed on the conductor as a uniform close fitting tube of insulation.

The idea of using spongy or foamed hydrocarbons as conductor insulation is not new. A British patent[3] issued in 1930 contemplates such a structure, and numerous United States patents of more recent dates cover various aspects of cellular hydrocarbon insulation. The problem is one of forming the cylinder of aerated plastic in an extrusion process operating at high speed, and producing a closely controlled uniform covering having precise physical and electrical properties. The original development work was carried on by F. P. Lyons of the Western Electric Company in co-operation with the author. This early work demonstrated that material having the desired range of properties could be applied in a continuous extrusion process, and subsequent work has shown that the necessary control of properties and speed of extrusion can be achieved.

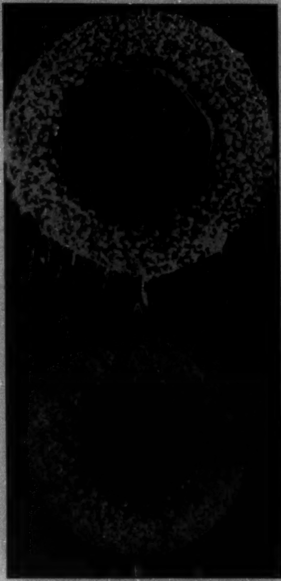


Fig. 2. Cross section of expanded-polyethylene insulation from 19-gauge conductor (magnified 75 times)

A. 35-per-cent air
B. 55-per-cent air

The expansion of polyethylene is accomplished by methods similar to those employed in the making of many of the numerous polymer and rubber "foams." The process used to produce the cellular polyethylene involves the addition at the extruder of a chemical blowing agent which decomposes under heat and releases nitrogen gas. By proper mixing and process control this nitrogen gas can be entrapped in the polyethylene in the form of very small discrete bubbles, thus achieving the cellular structure shown in Fig. 2. It is interesting to note that the foamed plastic tends to form a desirable "skin" of solid material on the inner surface over the conductor Fig. 2(A).

Various degrees of expansion can be achieved as required, by varying the amount of blowing agent and by other means. The degree of expansion, or per-cent entrapped gas, can be determined readily by weighing a sample of insulation, with the conductor removed, on an analytical balance. The inside and outside diameter of the cylinder of insulation and the density of solid polyethylene are required to complete the determination.

The composite dielectric constant obviously varies with the degree of expansion. In a coaxial configuration, this effect is calculable from the formula for the dielectric constant of a mixture,[4] the relation being given by the following

$$\frac{\epsilon-\epsilon_p}{3\epsilon}=V\frac{(\epsilon_a-\epsilon_p)}{(\epsilon_a+2\epsilon)}$$

where

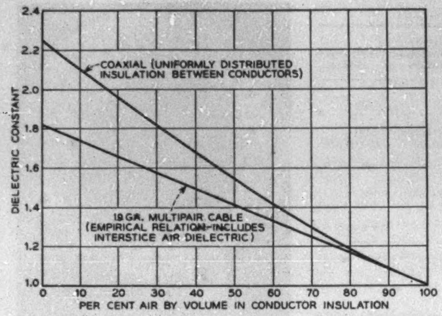


Fig. 3 (left). Change in dielectric constant of coaxial cable and multipair cable with degree of expansion of polyethylene conductor insulation

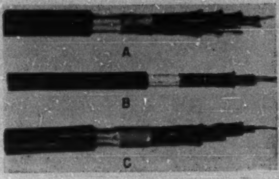


Fig. 5. Polyethylene-insulated cables—alpeth sheath

A. Grandville—Zeeland (Mich.) project: 51 pairs; 19-gauge conductors; expanded-polyethylene insulation

B. Cooperstown—Cherry Valley (N. Y.) project: 26 pairs; 19-gauge conductors; solid-polyethylene insulation

C. Trout Lake—St. Ignace (Mich.) project: 51 pairs; 19-gauge conductors; solid-polyethylene insulation

ϵ = composite dielectric constant
ϵ_p = dielectric constant of polyethylene = 2.26
ϵ_a = dielectric constant of added material, in this case (air) = 1
V = volume fraction = $\dfrac{\text{per-cent air}}{100}$

The upper curve in Fig. 3 was determined in this manner. However, in multipair cable, the dielectric constant, is not amenable to calculation and must be determined experimentally.[5] The empirical relation between the per-cent air in the insulation and the dielectric constant in the cable for a typical design is shown in the lower curve in Fig. 3. Effective dielectric constants as low as 1.40 have been achieved in expanded polyethylene cables.

Solid-polyethylene-insulated cables are more costly for given transmission characteristics than those insulated with paper or pulp and, therefore, are restricted to special uses, where a system saving can be obtained in spite of the higher first cost. One such use is for small aerial toll cables in rural areas where, with paper-insulated cable, maintenance costs are likely to be high. There are several factors which tend to increase the maintenance costs in such cables. For instance, in small isolated cables, lightning troubles are common because of the high sheath resistance. While the incidence of sheath breaks from other causes is usually neither more nor less than in other aerial cables, maintenance is more difficult and costly because of inaccessibility. Maintaining gas pressure on these cables is expensive for the same reason.

In the case of a sheath break, with no insulation damage in polyethylene-insulated cables, there is no interruption in service due to entrance of moisture into the cable, and repairs can be made on a routine maintenance basis. Gas-pressure maintenance is unnecessary and its omission effects appreciable annual savings. Because these cables usually contain not more than 51 pairs, the first-cost penalty for the use of solid polyethylene, in terms of cents per foot, is small. Small aerial toll cables in rural areas, therefore, are the most promising candidates for solid-polyethylene insulation.

The solid-polyethylene insulation development has progressed to the point where several cables have been made for field trials. The first of these was at Cooperstown, N. Y., where approximately 4 miles of 26-pair 19-gauge cable was installed aerially in 1950. The sheath on this cable was a composite of aluminum and polyethylene commonly known as "alpeth" sheath. The Cooperstown cable, see Fig. 4, was one of the first to be installed by the prelashing method, which has been developed as a means for effecting economies in placing aerial cable.

A second trial installation of solid-polyethylene-insulated cable was made in 1951 between Trout Lake and St. Ignace, Mich., a distance of 28 miles. This cable contained 51 pairs of 19-gauge conductor, and was also covered with alpeth sheath. Since that time, solid-polyethylene cables have been installed in other locations where the anticipated maintenance savings were believed to justify the higher first cost.

Development of expanded-polyethylene insulation has been carried on concurrently with that of solid polyethylene. In expanded-polyethylene insulation the cost, as would be expected, varies with the degree of expansion. There are two reasons for this:

1. As the proportion of gas is increased, less polyethylene is used in the insulation.

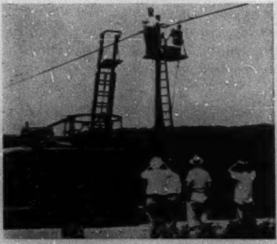


Fig. 4. Installation of Cooperstown—Cherry Valley (N. Y.) cable

Table II. Comparative Data on 51-Pair 19-Gauge Cables

Type of Insulation	Inside Diameter of Sheath, Inches	Mutual Capacitance, Microfarads per Mile	Attenuation, Decibels (Db) per Mile (70 F)* 1 Kc	60 Kc	150 Kc	Average Unbalance to Ground, Micromicrofarads per 1,500 feet	Minimum Dielectric Strength, Kv Conductor to Conductor	Core to Sheath
Solid polyethylene (Trout Lake project)	0.92	0.082	1.24	4.51	6.86	104	10	10
Paper (standard CNB)	0.80	0.084	1.26	4.94	7.79	310	0.7	1.4
Expanded polyethylene (Grandville project)	0.97	0.066	1.11	3.92	5.96	118	3	10
Paper (standard DNB)	0.94	0.066	1.11	4.05	6.40	240	0.7	1.4

* Computed from primary constants.

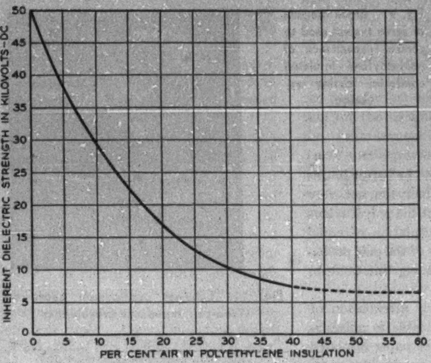


Fig. 6. Inherent dielectric strength of expanded polyethylene versus degree of expansion; 19-gauge 0.064-inch diameter over dielectric conductors; short samples immersed in water

2. Because of the lower dielectric constant, the cable can be made smaller for the same attenuation, resulting in savings in sheathing materials.

An initial trial installation of about 19 miles of 51-pair 19-gauge expanded-polyethylene-insulated cable has recently been completed between Grandville and Zeeland, Mich. This route is to be developed for *N*-carrier, and since recent Bell Systems studies have indicated that lower overall costs will result if low-attenuation cable is used, the Grandville-Zeeland cable was designed for a capacitance of 0.066 microfarad per mile, rather than the 0.084-microfarad-per-mile capacitance for which earlier polyethylene-insulated cable were designed. The Cooperstown, Trout Lake, and Grandville cables are shown in Fig. 5.

It is of interest to compare polyethylene-insulated cables with paper-insulated cables having the same voice-frequency attenuation. Some of the more important characteristics are shown in Table III. It will be noted that the polyethylene-insulated cables excel in dielectric strength. The values shown in Table II are the test voltages that all the reels of cable were required to withstand between conductors, and from core to sheath. These voltages are naturally made lower than the inherent dielectric strength of the insulation to allow for minor manufacturing irregularities and thus avoid excessive rejections. Some idea of the inherent dielectric strength of solid- and expanded-polyethylene insulation can be obtained from tests on short samples of insulated conductor immersed in water. In Fig. 6, the inherent dielectric strength obtained in this manner for conductors with a 14-mil wall is plotted against per-cent air in the insulation. It will be noted that the dielectric strength falls off rapidly as the polyethylene is expanded.

The capacitive unbalance to ground, or difference in direct capacitances of the wires of a pair to ground, is a rough indication of one important factor in the susceptibility of cable circuit to noise and interference. The electric disturbances which cause noise in the cable may come from atmospheric-static, radio-station, power line, or telephone-plant sources. The first three sources energize the cable circuits via the surface transfer impedance[6] of the sheath, or by way of an open-wire tap; the fourth, via all of the conductors in the cable. For this reason the two wires of a pair should have nearly equal capacitance to the surrounding pairs and to the sheath. To achieve this condition the cylinders of insulation on the two wires of a pair must be alike in size and dielectric constant. The low value of capacitive unbalance obtained with solid-polyethylene insulation is a result of the remarkable uniformity with which this insulation is extruded. The value of 104 micromicrofarads per 1,500 feet for the capacitive unbalance to ground is on the average only 0.4 per cent of the direct capacitance of either wire to ground.

Obtaining nearly equal capacitance to ground with expanded polyethylene is somewhat more difficult than with solid polyethylene since, with the former, the degree of expansion (i.e., dielectric constant) may vary as well as diameter. The recent development of a device which continuously measures the capacitance of the conductor to the cooling water during extrusion of the insulation has been of considerable help in maintaining low capacitive unbalance to ground in expanded-polyethylene cables. Space will not permit a detailed description of this device. However, the device, known as a capacitance monitor, consists essentially of a hybrid bridge with a guarded electrode on one side and a standard capacitor on the other. The electrode operates submerged in the cooling-water trough

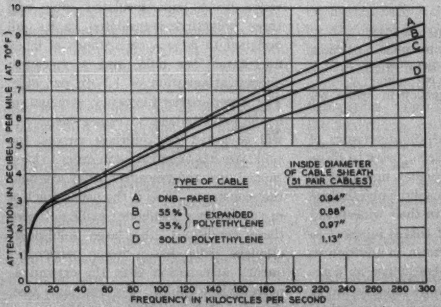


Fig. 7. Attenuation versus frequency; 19-gauge cables having 0.066-microfarad-per-mile capacitance

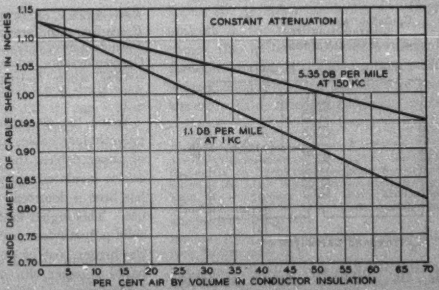


Fig. 8. Change in cable diameter with degree of expansion of polyethylene conductor insulation; 19-gauge 51-pair cable

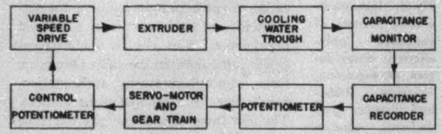


Fig. 9. Block diagram of servo system used to control capacitance of polyethylene-insulated conductor during extrusion

of the extruder. The capacitance of a controlled length of insulated conductor can be balanced by the proper setting of the standard capacitor. If the standard capacitor is set at the desired capacitance of the insulated conductor, and the output of the bridge fed to a conventional recorder, a continuous record of the deviations from the nominal capacitance is obtained.

Going one step further, the output of the bridge can be used to control the speed of the extruder capstan through an appropriate servo system. The extruder speed, in turn, will determine the conductor diameter, so that the diameter can be varied to compensate for uncontrollable changes in the degree of expansion, and thus maintain the capacitance constant. The coaxial capacitance to water and the capacitance to ground in the cable are not the same, but they are related in a linear fashion. Control of one is then effectively a control of the other. A block diagram of the servo system is shown in Fig. 9, and a schematic of the capacitance monitor in Fig. 10. In a recent experimental cable made with servo control of capacitance, the unbalance to ground was only 35.2 micromicrofarads per 1,500 feet, or 0.18 per cent of the capacitance of one conductor to ground in the cable.

The other important factor in unbalance is the uniformity of the twisting, that is, the extent to which wire and mate form symmetrical helixes around the center line of the pair. Polyethylene-insulated pairs appear to be better in this respect, for reasons which are not very obvious. The fact that the polyethylene forms firm tubes of insulation of equal size on both conductors of the pair probably is a factor in achieving this uniform twisting.

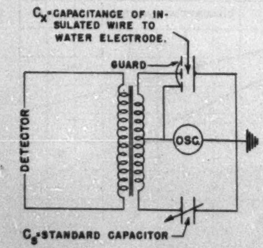


Fig. 10. Schematic diagram of capacitance-sensing unit or monitor, used at extruder for the continuous measurement of conductor capacitance during extrusion

The carrier-frequency attenuation of polyethylene-insulated cable is substantially lower than that of paper-insulated cable having approximately the same voice-frequency attenuation. In the case of solid-polyethylene cable there are two reasons for this lower carrier-frequency attenuation:

1. The inductance is higher and the conductance is lower than those in comparable paper-insulated cable. The higher inductance is a result of the greater separation between the wires of a pair and the lower conductance is a result of the lower power factor of polyethylene compared to that of paper.
2. The voice-frequency attenuation is relatively independent of inductance and conductance.

The approximate voice-frequency and carrier-frequency attenuation constants may be calculated by means of the usual formulas, namely

Voice frequency

$$\alpha = 8.686\sqrt{\frac{RC\omega}{2}}\text{ db per mile}$$

Carrier frequency

$$\alpha = 4.343\left[R\sqrt{\frac{C}{L}} + G\sqrt{\frac{L}{C}}\right]\text{ db per mile}$$

where

R = resistance, ohms per mile
C = capacitance, farads per mile
L = inductance, henrys per mile
G = conductance, mhos per mile
$\omega = 2\pi f$, where f = frequency in cycles per second

While the electrical characteristics of polyethylene cables are superior to those of paper cables, the higher first cost of cables insulated with solid polyethylene has been a deterrent to their widespread use. This higher first cost is inherent in solid polyethylene because, in addition to the higher cost of polyethylene as compared to paper, the cables must be larger for the same voice-frequency attenuation. It will be noted that the 51-pair Trout Lake-St. Ignace cable is 15 per cent larger in diameter than the comparable paper cable. The larger size is necessary because of the higher effective dielectric constant, which is approximately 1.80 in solid polyethylene cable as compared to 1.60 in a typical paper cable. The effective dielectric constant is higher in the case of solid-polyethylene insulation, because of the lesser amount of air space which can be incorporated in the dielectric between wires.

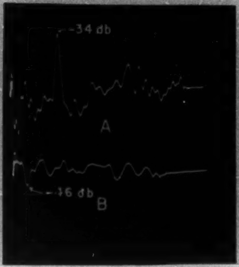


Fig. 11. Echo-set oscillograph traces of video-pair impedance irregularities

A. 2,700-foot-length ribbon-and-string video pair (maximum echo = 34 db)
B. 2,000-foot-length expanded-polyethylene video pair (maximum echo = 46 db)
Note: Calibration attenuation was on the same setting for both exposures

As was shown in Fig. 3, the dielectric constant can be decreased by expanding the polyethylene. A value as low as 1.40 has been attained experimentally. The effective dielectric constant of the Grandville-Zeeland cable was 1.54; it was obtained by expanding the conductor insulation to the point where 35 per cent of the volume was gas.

The curves of Fig. 7 represent the attenuation over the *N*-carrier frequency range for a paper-insulated cable and three polyethylene-insulated cables, all designed to have a capacitance of 0.066 microfarad per mile, and a voice-frequency attenuation of 1.1 db per mile. The lower carrier-frequency attenuation of the polyethylene cables is evident. It will be noted that in these curves, the cable size has been varied in order to keep capacitance and voice-frequency attenuation constant. This is the usual practice in designing telephone cable. Under these conditions, the large solid-polyethylene cable has a lower carrier-frequency attenuation than the expanded-polyethylene cables. For the same size, the expanded-polyethylene cables naturally have lower attenuation at both voice and carrier frequencies. The size of the

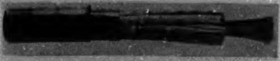


Fig. 12. Cable containing eight expanded-polyethylene insulated video pairs and 51 pairs of 22 gauge, also insulated with expanded polyethylene

polyethylene cables varies with the degree of expansion, as shown in Fig. 8.

Since the dielectric strength decreases as the degree of expansion increases, the saving in first cost must be balanced against the value of the reduction in reliability. As mentioned before, the Grandville-Zeeland project was the first commercial installation of expanded-polyethylene cable. A very moderate degree of expansion was chosen for this project. However, as more experience with expanded polyethylene is gained, it should be possible, for a given application, to determine the degree of expansion which strikes the optimum balance between dielectric strength and dielectric constant, giving proper weight to the mechanical properties and to cost factors.

The expanded polyethylene is also being developed for the 16-gauge video pairs which are used for local television transmission. In this case, the problem is one of impedance uniformity. The standard ribbon and string design illustrated in Fig. 1 has not been entirely satisfactory in this respect. The impedance deviations between cable lengths and internal impedance irregularities are reduced by the use of expanded-polyethylene insulation on the conductors. Fig. 11 shows the improvement in internal echoes. These oscillograms show the Echo Set[7] oscillograph traces for an experimental expanded-polyethylene video pair and a standard ribbon-and-string-insulated video pair. The maximum echoes in the former are about 46 db down from the measuring pulse, as compared to about 34 db for the latter.

A field trial of the new video pair, consisting of 7,000 feet of cable containing eight video pairs and 51 22-gauge pairs was installed near Tulsa, Okla., in 1953. The 22-gauge pairs, which are intended for local telephone use, were also insulated with expanded polyethylene. This cable is shown in Fig. 12. The use of expanded polyethylene for the finer-gauge exchange pairs has not yet been evaluated. In the final analysis, it will probably depend largely on cost factors.

In conclusion, the status of the development may be stated briefly as follows: The solid-polyethylene insulation has been introduced into the telephone plant for restricted use. The expanded-polyethylene insulation has reached the field trial stage for 19-gauge toll cable, 22-gauge exchange circuits, and broadband video pairs.

References

1. MODERN PLASTICS ENCYCLOPEDIA (book). Plastic Catalog Corporation, New York, N. Y., 1950.

2. REPORT OF CONFERENCE ON RADIO-FREQUENCY CABLES. PART II—GENERAL CHARACTERISTICS OF POLYETHYLENE, J. W. Shackleton. *AIEE Transactions*, vol. 64, Dec. 1945, pp. 912–16.

3. British Patent *No. 334,840*, 1930.

4. THE DIELECTRIC CONSTANT OF CRYSTALLINE POWDERS, C. J. F. Boettcher. *Recueil des Traveaux Chimiques des Pays-Bas*, The Hague, Netherlands, vol. 64, 1945, p. 47.

5. MEASUREMENTS IN MULTIPAIR CABLES, J. T. Maupin. *Bell System Technical Journal*, New York, N. Y., vol. 30, July 1951, pp. 652–67.

6. ELECTROMAGNETIC THEORY OF COAXIAL TRANSMISSION LINES AND CYLINDRICAL SHIELDS, S. A. Schelkunoff. *Bell System Technical Journal*, New York, N. Y., vol. 13, Oct. 1934, p. 532.

7. PULSE ECHO MEASUREMENTS ON TELEPHONE AND TELEVISION FACILITIES, L. G. Abraham, A. W. Lebert, J. B. Maggio, J. T. Schott. *AIEE Transactions*, vol. 66, 1947, pp. 541–48.

8. BELL SYSTEM CABLE SHEATH PROBLEMS AND DESIGNS, F. W. Horn, R. B. Ramsey. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1811–16.

◆

No Discussion

Flux Resetting Characteristics of Several Magnetic Materials

HOOBERT HUHTA
ASSOCIATE MEMBER AIEE

EXPERIMENTAL data are presented on the rate of change of flux in four materials used in magnetic amplifiers. A test circuit which applied a constant magnetizing force to a test core was employed.[1] The rate of change of flux was then observed as the induced voltage in a pickup coil. By observing the induced voltage wave on an oscilloscope, one can determine the peak and average rate of change of flux, along with the wave shape as a function of time. This information is helpful in understanding and improving the performance of magnetic amplifiers, as will be discussed later.

The recent trend in magnetic amplifier design has been toward the use of a d-c or an a-c hysteresis loop as a design basis.[2,3] Unfortunately, the shape of the operating hysteresis loop is partially determined by the external circuitry.[4] In a self-saturating magnetic amplifier with high control circuit impedance the flux is reset with approximately a constant magnetizing force. This brief study deals with the flux change under the influence of a constant external magnetizing force. The results are qualitatively applicable directly to the high control circuit impedance magnetic amplifier.

Test Methods

All the materials tested were in the form of either 1/4- or 1/2-inch-wide tape, wound to form a gapless toroidal core. These cores had a nominal $1\frac{1}{2}$-inch inside diameter and about 1/4 inch of build-up. Each core had two 1,000-turn excitation windings and a third winding of about 100 turns for instrumentation purposes.

Fig. 1 shows the circuit used to excite the test cores. The primary function of this circuit is to allow a constant current in winding *B* of the test core to drive the flux from saturation in one direction to saturation in the opposite direction without allowing current to flow in winding *A*. When the tyratron conducts, a voltage appears across resistor R_1 which overcomes the battery and drives current through winding *A*. This voltage pulse is large enough to saturate the core and has a repetition rate adjustable in 1-cycle

---

Paper 54-161, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 19, 1953; made available for printing December 4, 1954.

HOOBERT HUHTA is with the General Electric Company, Schenectady, N. Y.

The material for this paper was collected under the direction of H. W. Lord while the author was on an assignment from the Advanced Engineering Program to the Research Laboratory of the General Electric Company.

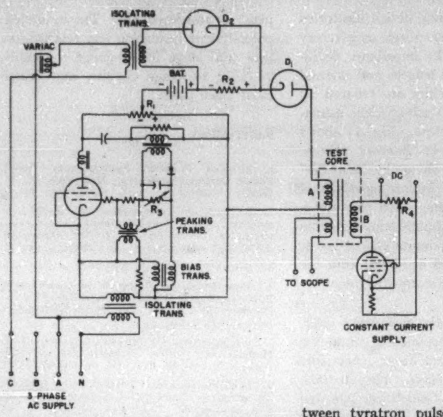


Fig. 1 (left). Test circuit

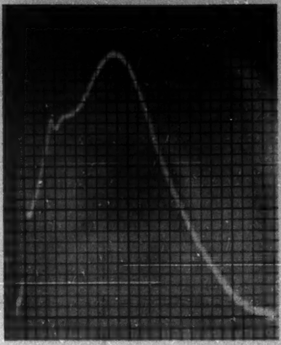


Fig. 2 (right). Rate of change of flux versus time curve for a 65-Permalloy core of 0.005 - inch - thick tape resetting in 0.044 second

steps by means of resistor R_3. During the nonconducting half-cycle of the tyratron, diode D_2 conducts to place a high voltage across R_2, hence a high reverse voltage on diode D_1. The supply voltage for D_2 leads the tyratron supply voltage by 30 degrees to cut off the current in diode D_1 more abruptly. When neither the tyratron nor diode D_2 are conducting, the battery voltage keeps D_1 from conducting. Winding B is supplied with essentially a constant current from the pentode supply at all times. A third test core winding is connected to a cathode-ray oscilloscope for observing the induced voltage, i.e., the rate of change of flux.

The data presented here were obtained in the following manner. The circuit as shown was supplied from a commercial 60-cycle line. After a test core was placed in the circuit, R_1 was adjusted to assure that the core would saturate whenever the tyratron conducted. The battery and the voltage across R_2 were made large enough to prevent conduction of diode D_1, unless the tyratron was conducting. R_3 was then adjusted to allow the tyratron to conduct every half cycle. The current in winding B was then varied by means of R_4 until the flux would just reset in the nonconducting half-cycle of the tyratron. This condition can be found by observing on the oscilloscope that the rate of change of flux just reaches zero as the voltage from the tyratron supply begins to set the flux for the next cycle. The oscilloscope trace was photographed during the resetting portion of the cycle to record the rate of change of flux as a function of time. Resistor R_3 was readjusted to change the period between tyratron pulses. Data were recorded as before. Fig. 2 illustrates the oscilloscope face showing a typical rate of change of flux versus time curve.

Results

Fig. 3 is a plot of the average rate of change of flux density as a function of the external magnetizing force minus the d-c coercive force, hereafter referred to as the excess magnetizing force. The average rate of change of flux was found by taking twice the d-c saturation flux density and dividing by the time required to reset the flux. The excess magnetizing force was calculated by converting the current in winding B to units of magnetomotive force and subtracting the coercive force found on a d-c hysteresis loop. These data are not intended as a reference, since the results vary noticeably between core samples, but they do indicate the general trends. One observes that for a given material and excess magnetizing force the thicker laminations reset more slowly. It is also noteworthy that magnetic annealing causes Deltamax to reset more slowly.

The shapes of the rate of change of flux versus time curves show interesting variations. Fig. 4 is a plot of such curves for three Deltamax cores of varying lamination thickness. The actual photographs were replotted to facilitate superimposing the results. Notice the increased tendency toward a second hump as the thickness increases.

Two rather different shapes of rate of change of flux versus time curves are shown by specimens of Perminvar and Supermalloy in Fig. 5. Supermalloy has

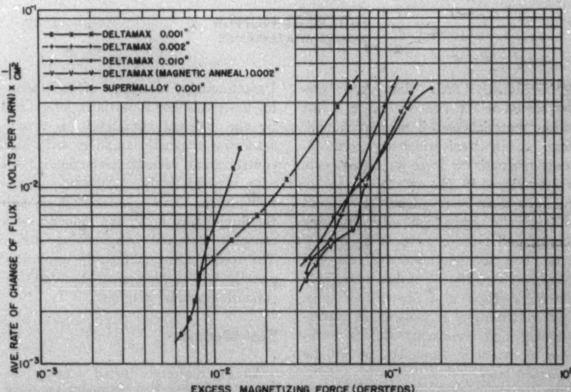


Fig. 3. Average rate of change of flux density versus excess magnetizing force for several materials

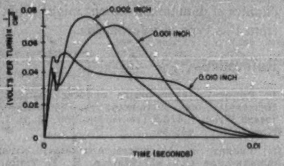


Fig. 4. Rate of change of flux versus time for Deltamax

a very high rate of change of flux early in time. The exponential effect is probably caused by the sloping sides of the hysteresis loop decreasing the effective excess magnetizing force as the flux approaches negative saturation. Perminvar has the double hump typical of thick or low resistivity materials but it also has a small unexplained notch on the rising part of the curve.

All the rate of change of flux curves show a pronounced spike at the immediate start of the curve. This high rate of change of flux corresponds to the flux change from saturation to a point below the outer knee of the magnetization curve. Presumably, this is the energy stored in air-cored inductance and domain rotation.

Fig. 6 is a rate of change of flux curve taken at a very low value of excess magnetizing force. Besides being very irregular this same curve is not retraced on successive resetting cycles. The core sample shown exhibited the greatest irregularities, but all the materials tested gave similar results.

All of the data presented so far were taken with the flux resetting from saturation. When the flux is not reset from saturation the results are very different. The curves in Fig. 7 show the rate of change of flux as a function of flux density with the external magnetizing force held constant. The core was magnetized to some flux density, given by the intersection of the rate of change of flux curve and the flux density axis, and allowed to reset to negative saturation. All the curves that start from a flux density less than saturation have both a peak and an average rate of change of flux which are greater than that for a curve starting from saturation. This suggests that it takes an appreciable amount of time either to establish domain boundaries or to start them moving, once they have been subjected to a high magnetizing force.

These results do not offer an answer to the question of why magnetic materials behave as they do. However, it is hoped that the technique applied here will be useful in deriving more information leading to the better understanding of ferromagnetic materials and their operation in magnetic amplifiers.

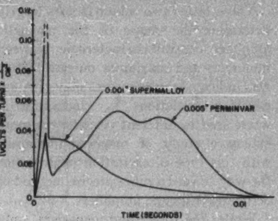


Fig. 5. Rate of change of flux versus time for Perminvar and Supermalloy

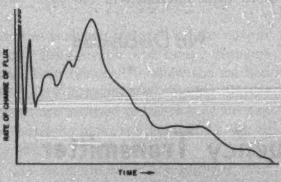


Fig. 6. Curve of rate of change of flux versus time at low magnetizing force. Perminvar core of 0.005-inch tape

Application to Magnetic Amplifiers

In a qualitative way results from this type of test can be applied immediately to high-control circuit-impedance self-saturating magnetic amplifiers. For example, the effect of the flux resetting characteristic upon amplifier gain can be studied.

In such an amplifier the core presumably is subjected to a constant magnetizing force during the nonconducting half-cycle of supply voltage. At first glance, core materials which require the lowest excess magnetizing force to reset their flux in a time equivalent to 1/2 cycle of the supply voltage might be expected to give the highest gain. This is usually the most important factor but not the only criterion.

In the normal half-wave magnetic amplifier circuit shown in Fig. 8 the available reverse voltage on the gate winding is not sufficient to prevent conduction of the diode during the resetting period when the resetting period approaches times as short as 1/2 cycle. The resulting current causes a loss in the flux-resetting ability as shown by the dotted area in Fig. 9.

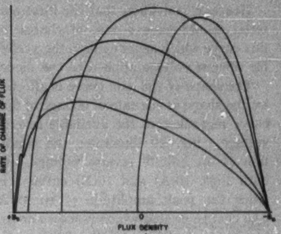


Fig. 7. Rate of change of flux versus flux density taken on a magnetically annealed Deltamax core of 0.002-inch tape

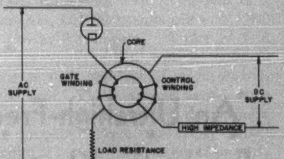


Fig. 8. Half-wave magnetic amplifier circuit

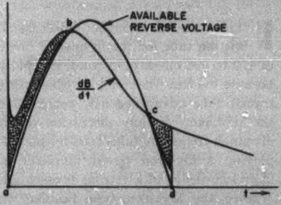


Fig. 9. The voltage resulting from flux resetting plotted with the available reverse voltage for a half-wave magnetic amplifier circuit

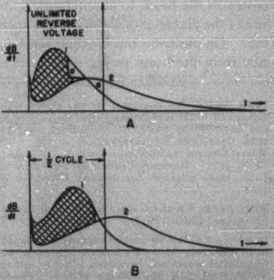


Fig. 10. Variation in the amount of flux reset caused by the shape of the rate of change of flux versus time curve

Hence, it is evident that the ideal material should have a rate of change of flux curve which never exceeds the available reverse voltage applied to the diode.

The shape of the gain curve is affected by the shape of the rate of change of flux curve regardless of the available reverse voltage. Fig. 10 illustrates the condition of an infinite reverse voltage. In both Figs. 10(A) and 10(B) curve 1 is twice the peak amplitude of curve 2. The change in area which is shown cross-hatched is greater for 10(B). This means that the material of 10(B) would have a higher gain, assuming the change in control current were the same in both cases.

These latter two minor factors largely determine the shape of the magnetic amplifier control characteristic between minimum and maximum output. They also indicate why two cores may give unexpected variations in characteristics when used in different circuits. A more extensive study of magnetic materials with this type of excitation might point the way toward a more precise method of designing magnetic amplifiers.

While the preceding considers only magnetic amplifier aspects, this type of approach may also be useful in core materials evaluation and development. Even those interested in the basic mechanics of domain movement might profit from this type of information.

References

1. PULSE CHARACTERISTICS OF RECTANGULAR-HYSTERESIS-LOOP FERROMAGNETIC MATERIALS, **Joseph Wylen**. *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 648–56.

2. THEORY OF MAGNETIC AMPLIFIERS WITH SQUARE-LOOP CORE MATERIALS, **H. F. Storm**. *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 629–40.

3. PREDETERMINATION OF CONTROL CHARACTERISTICS OF HALF-WAVE SELF-SATURATED MAGNETIC AMPLIFIERS, **Henry Lehman**. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 2097–2103.

4. DYNAMIC HYSTERESIS LOOPS OF SEVERAL CORE MATERIALS EMPLOYED IN MAGNETIC AMPLIFIERS, **H. W. Lord**. *AIEE Transactions*, vol. 72, pt. I, March 1953, pp. 85–88.

No Discussion

An Ultrahigh-Frequency Transmitter Employing Klystron Power Amplifiers

W. H. SAYER
NONMEMBER AIEE

IN THE search for a higher power transmitting tube for the frequency range of 470 to 890 megacycles (mc), it rapidly became evident that only a tube using transit time advantageously could solve the problem. In any discussion of an ultrahigh-frequency (uhf) transmitter employing a klystron power amplifier, a basic knowledge of klystrons is essential.

Since the klystron type familiar to many is the small low-powered reflex klystron oscillator, a review of the high-powered cascade klystron amplifier is given.

The electron gun of the klystron concentrates the beam to approximately the size of the diameter of the drift tube opening. The electrons receive their acceleration from the beam power supply. The electron beam enters the drift tube, then proceeds down the drift tube, and is kept from spreading by the axial magnetic field. When the electrons reach the first gap, they are either accelerated further or decelerated by the radio frequency voltage set up by the first tuned circuit. By the time these electrons are at the second gap, they have had a chance to form discrete bunches. The electron bunches excite the second tuned circuit improving the bunching and increasing the gain of the klystron. The electron beam is in the form of bunches by the time of arrival at the output gap. The output gap must present the correct impedance to transfer effectively all the radio frequency power from the final bunched beam to the loaded cavity. The collector then dissipates the remaining power of the electron beam and is usually water-cooled.

This general type of tube has been used previously in wide band operation in France[1] and the United States. Because the early tubes have cavities in the vacuum system, their method of tuning has been to change the gap spacing in the drift tubes. The only commercial tubes available with demountable cavities outside the vacuum system are the Eimac klystron series, type *3K20, 000LA, LF, LK*. The demountable cavities allow considerable tuning range (720 to 890 mc on the *3K20, 000LK*) without changing the gap spacing, and allow the designer to have control of the external tuned circuits for his application.

Fig. 1 shows the complete 5-kw Du Mont transmitter with the front doors open, designed for use with the klystrons. The two center frames are for the exciter and modulated amplifier. The sound channel klystron and power supply are on the left, and the visual channel klystron and power supply on the right. The high-voltage power transformers, variable-voltage transformers, side-band filter, diplexer, and heat exchanger are installed externally to the frames. Fig. 2 is a block diagram of the system. The frequency control for both visual and aural transmitters is shown in Fig. 3. The frequency modulation of the aural

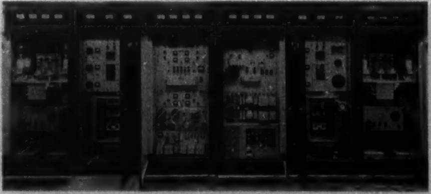


Fig. 1 (right). The complete transmitter

Paper 54-166, recommended by the AIEE Television and Aural Broadcasting Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted September 23, 1953; made available for printing December 23, 1953.

W. H. SAYER is with the Allen B. Du Mont Laboratories, Inc., Clifton, N. J.

In a project of this size, credit cannot be given here to all who made the work practical. However, this transmitter would not have been possible without W. W. Eitel, J. A. McCullough, and the Research Division at Eitel-McCullough, Inc. (Eimac), San Bruno, Calif. We should like to express our sincere appreciation to them for their help and inspiration, as well as to all our associates of the Allen B. Du Mont Laboratories, Inc.

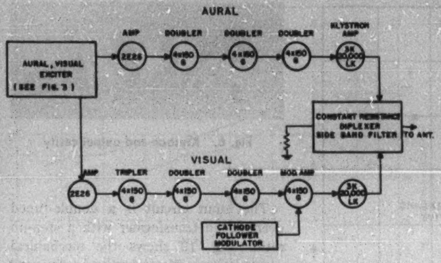

AURAL
AURAL, VISUAL EXCITER (SEE FIG. 3)
AMP
2E26
DOUBLER
4x150 G
DOUBLER
4x150 G
DOUBLER
4x150 G
KLYSTRON AMP
3K 20,000 LK
CONSTANT RESISTANCE DIPLEXER SIDE BAND FILTER
TO ANT.
VISUAL
AMP
2E26
TRIPLER
4x150 G
DOUBLER
4x150 G
DOUBLER
4x150 G
MOD. AMP
4x150 G
3K 20,000 LK
CATHODE FOLLOWER MODULATOR

Fig. 2. Block diagram of system

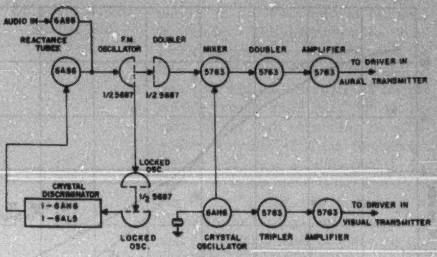

AUDIO IN
6AS6
REACTANCE TUBES
6AS6
F.M. OSCILLATOR
DOUBLER
1/2 5687
1/2 5687
MIXER
5763
DOUBLER
5763
AMPLIFIER
5763
TO DRIVER IN AURAL TRANSMITTER
LOCKED OSC
CRYSTAL DISCRIMINATOR
1-6AH6
1-6AL5
1/2 5687
LOCKED OSC.
6AH6
CRYSTAL OSCILLATOR
5763
TRIPLER
5763
AMPLIFIER
TO DRIVER IN VISUAL TRANSMITTER

Fig. 3. Aural and visual exciter

oscillator is produced by a reactance tube. A second reactance tube controls the average frequency of the oscillator and the control voltage, for this tube is derived from a crystal discriminator.[3] Two dividers locked to the frequency-modulation oscillator drive the crystal discriminator. The stabilized frequency-modulated oscillator output is doubled, then mixed with the visual crystal oscillator frequency. The exciter output frequencies are 91.718 mc for sound and 60.77 mc for visual for a channel-57 transmitter. The aural exciter frequency is then multiplied by 8 with three doublers using *4X150G*'s and the visual exciter frequency is multiplied by 12 with one tripler and two doublers.

The visual overtone crystal operates at approximately 20.25 mc, and any drift in this crystal oscillator frequency also changes the aural frequency. However, the difference (4.5 mc) is dependent upon the stable crystal discriminator operating at a low frequency.

The final visual frequency doubler drives the modulated amplifier which is shown in Fig. 4. This circuit is a grounded grid type as far as radio frequency is concerned but is grid-modulated with video. The plate circuit is a wave guide resonator with capacity output coupling. The video modulation from a low-impedance cathode-follower modulator is applied to the grid through a small radio-frequency (r-f) choke. The r-f drive is applied to the grid-cathode tuned circuit by a coupling loop fastened to the sliding short. Because the grid of the tube has a long electrical length at 730 mc, it is necessary to tune the grid to ground with a 1/2-wave line. This modulated amplifier is stable and with the effective grid to ground impedance slightly inductive, the linearity is improved considerably, as shown in Fig. 5.

The final linear amplifier is the Eimac klystron tube, type *3K20, 000LK* as shown in Fig. 6. One of the demountable cavities is shown fastened on the output gap. The movable end tuning plungers have been removed to show the method of cavity construction. Spring fingers which are fastened to the ring on the top of the cavity make contact to the 1/4-inch copper discs on the tube.

The basic circuit of the klystron used is shown in Fig. 7. The filament is a source of electrons and these electrons are accelerated to the cathode by the bombarder supply. This cathode is the source of electrons for the electron gun of the klystron and is made of tantalum. The tantalum button or cathode requires about 1,200 watts to emit electrons. The electrons are then accelerated to the anode in a modified pierce gun design by the high voltage beam supply. An auxiliary or prefocus magnetic circuit is used to center the beam in the drift tubes and to increase the over-all efficiency of the tube. After the beam of high-energy electrons enters the drift tube, a magnetic circuit keeps the electron beam from spreading. Any klystron amplifier tube must have an excellent vacuum to prevent gas focusing, as the operation of a gas focus tube is quite erratic and unpredictable. The magnetic field increases from about 110 gausses at the top pole to about 190 gausses at the bottom pole piece. This field is obtained by two large coils with 1,400 turns with a direct current of 2 amperes, and the bottom coil with 1,800 turns at 0.8 ampere. A plot of field versus length of drift tube is shown in Fig. 8, with positions of magnetic coils and gaps. The iron frame saves on d-c power because it completes the magnetic return path. The normal cavity current with the magnetic field adjusted properly is between 0.05 and 0.1 ampere. The cooling of the drift tubes requires a water flow of 1 gallon per minute, and the collector requires 8 to 10 gallons per minute.

The beam supply is a 3-phase full-wave bridge supply with a variable voltage transformer for voltage control up to 14,000 volts. This supply has a 2-section filter to reduce ripple resulting from line unbalance to prevent amplitude modulation of the beam. A single 3-phase full-wave bridge supply is used to get rid of spiking resulting from the start of mercury vapor rectifier conduction. The beam supply has a total current overload, and a cavity current overload in case of failure of the magnetic circuit. Both of these circuits are protected by thyrite elements as the so-called ground bus can approach beam voltage as a transient in

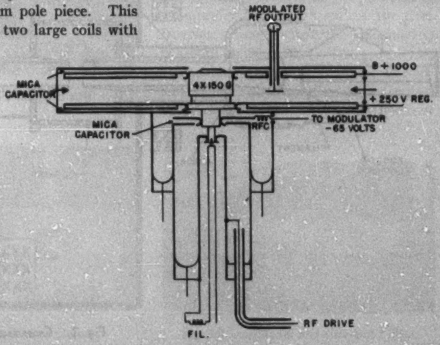

MODULATED RF OUTPUT
B+1000
4X150 G
MICA CAPACITOR
+250 V REG.
MICA CAPACITOR
RFC
TO MODULATOR -65 VOLTS
FIL.
RF DRIVE

Fig. 4 (right). Modulated amplifier

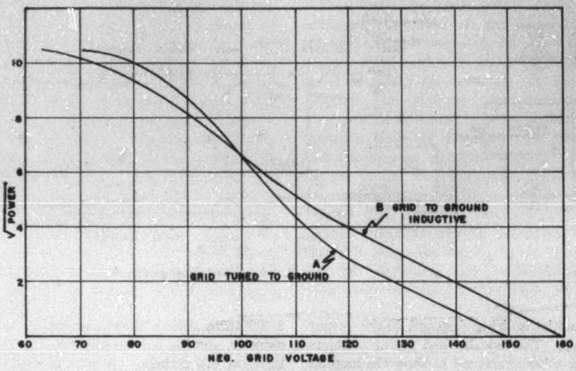


Fig. 5. Linearity of modulated amplifier

Fig. 6. Klystron and output cavity

case of short circuits at these voltage levels. A 37-ohm transient suppressor is also included in the high-voltage supply. The cathode is heated by a single-phase full-wave supply with a single-section filter delivering about 2,000 volts at 0.6 ampere. This bombarder supply has to be well insulated because the anode of the klystron is grounded, and therefore the supply is negative with respect to ground by the amount of the beam voltage. Fig. 9 shows the final assembly of the klystron with the cavities assembled and water plumbing assembled.

Having covered the mechanical aspects of the klystron, we are now concerned with using the klystron as a broadband amplifier. The input and center circuits are normally quite narrow and some thought must be given to broadbanding the amplifier.

The design of the first circuit in the klystron amplifier was made to accomplish maximum efficiency for a 5-mc bandwidth. In design of this type, two requirements must be met:

1. It is necessary to have a reasonably flat load on the upper side-band frequencies for correct operation of the video modulated amplifier.

2. It is necessary to have the maximum voltage across the first gap for velocity modulation of the electron beam for a given amount of drive power.

The input circuit is a double-tuned overcoupled transformer with a step-up ratio. Fig. 10 shows the mechanical construction. The inductive iris for controlling the coupling between the two tuned circuits is visible. The electron loading of the first gap was found to be quite small, allowing the main damping on the high-voltage side of the transformer to be a coupling loop terminated in a 50-ohm wattmeter. The voltage standing wave ratio at the input terminal was better than 1.2 to 1 from 1 mc below carrier to 4 mc above carrier frequency.

The center circuit on the middle gap must also have enough bandwidth. This circuit is a double-tuned overcoupled transformer similar to the input circuit. However, external damping is used on the low-voltage side of the transformer as the electron loading from the beam is small on the high-voltage side.

The output circuit is a single-tuned circuit with an adjustable coupling loop in the $3\frac{1}{8}$-inch coaxial line, shown in Fig. 11. The output circuit is broad enough for video use. In a theoretical

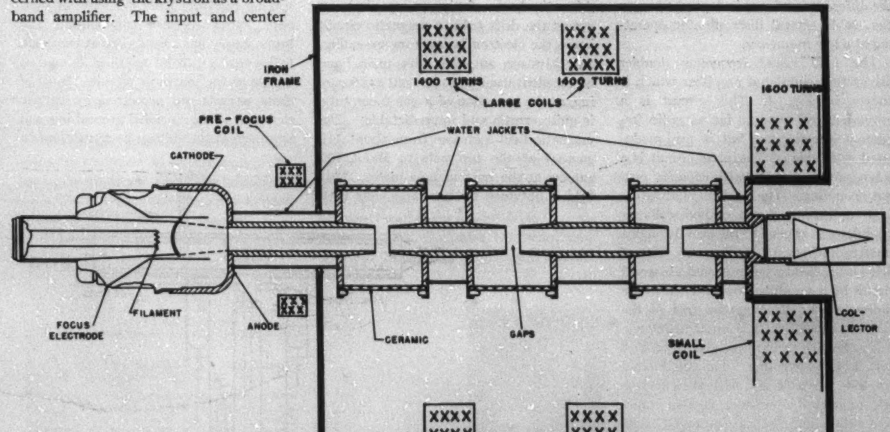


Fig. 7. Cross section of klystron

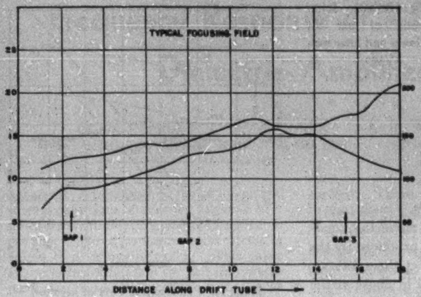


Fig. 8 (left). Typical focusing field

investigation,[3] it was shown that the output Q on the last tuned circuit for maximum output would be about 70 to 100 at 730 mc. Experimentally, the output Q needed for maximum output was 64, or a 3-decibel (db) bandwidth of 11.4 mc at 13-kv beam voltage and, as the beam voltage is increased, the loading must also be increased. This bandwidth is ample for video use, as a 4-mc side-band component would be down about 2 db.

To make the klystron amplifier flat over the passband, it is necessary to peak the center-tuned overcoupled circuit 2 db at the high end. This can be done because the average power in these high-frequency side bands is down about 20 to 30 db compared to the side-band components near the carrier. The 3.58-mc components, when transmitting National Television Systems Committee color, are down about 20 db.

With the klystron delivering full power (continuous wave) in broad-band operation, a power of 300 watts is dissipated in the wattmeter connected to the center circuit. The amplifier requires about 50 watts' total driving power under these conditions.

Circuit efficiency becomes more important in amplifiers used at ultrahigh frequencies. This determines the useful power output of the amplifier and the losses in the circuit. Correct design of by-passes and contact fingers in the uhf circuits become much more difficult because of the high circulating radio-frequency currents at these frequencies. Circuit efficiency is $100 \times Qu - Q_L/Qu$, where Qu is the unloaded Q and Q_L is loaded Q. Of course, it is necessary to couple to the tuned circuit to measure Q and a correction must be made for this coupling.

A Q measurement technique has been described[4] which corrects for the loading resulting from the coupling loop after the slotted line was used. This procedure is one of the most comprehensive on the measurement of unloaded Q of cavities. Use of the graphs in the report saves considerable time.

A comparison of circuit efficiency of a low-power tetrode and the klystron output circuit is as follows: At 730 megacycles

$$\text{circuit efficiency} = \frac{100(Qu - Q_L)}{Qu}$$

For a 4×150 tetrode

$$\text{efficiency} = \frac{300-80}{300} = 73 \text{ per cent}$$

output circuit≅4.7 micromicrofarads

For a klystron ampere

$$\frac{3{,}000-64}{3{,}000} = 97.9 \text{ per cent}$$

output circuit≅0.8 micromicrofarads

A typical tetrode is the *4X150G* used as a modulated amplifier to deliver 120-watt peak. The klystron can deliver 8 kw to the output cavity with only 175 watts lost in the output circuit including seals, ceramics, and copper losses.

The ceramic losses are less when the tuning plungers are closer to the ceramic windows near the high end of the tube frequency range where the unloaded Q is around 3,500.

The aural klystron amplifier is very similar, with the exception of the center-tuned circuit, which in this case does not

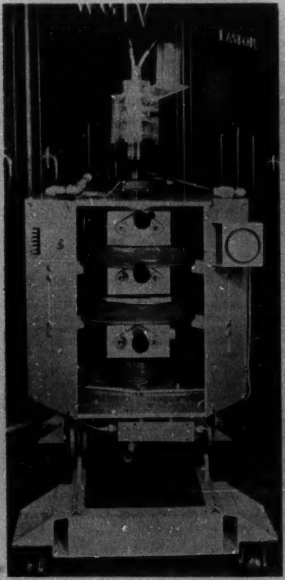


Fig. 9. Klystron assembly

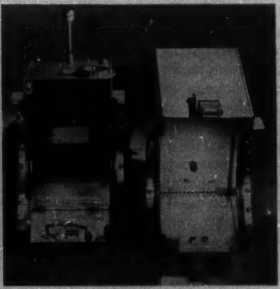


Fig. 10. Double-tuned circuit

Fig. 11. Output-tuned circuit

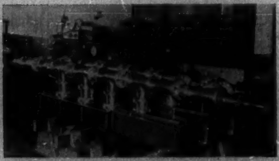


Fig. 12. Diplexer and side-band filter

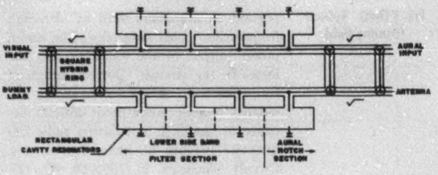


Fig. 13 (left). Schematic of uhf diplexer and filter network

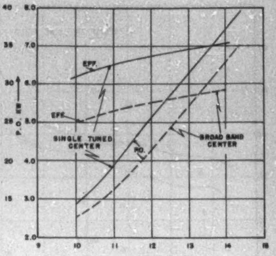


Fig. 15. Performance data of klystron

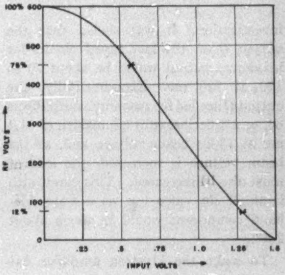


Fig. 16. Visual system linearity

have to be broad-banded, and therefore is a single-tuned cavity. The *Q* of the single-tuned circuit is much higher and the gain of the klystron is much greater. However, the driver in the aural chain supplies much more power than needed so the broad-band double-tuned input circuit is used. This simplifies the spare tube problem, as this can be pretuned for video and used in either transmitter amplifier. The center circuit is tuned on the high side of resonance to reduce the power gain still further which also gives a slight improvement in efficiency.

A side-band filter is necessary to get the sharp cutoff on the lower side band, and is combined with the diplexer. The diplexer is needed to combine picture and sound without interaction into a single-antenna feed line. This is shown in Fig. 12. The hybrid rings are 3/4-wave long on the arms and are shown schematically in Fig. 13. This combination[5] is a constant resistance device to present to the video transmitter a flat load at the sound rejection notches as well as at the side-band notches. The high *Q* cavities present short circuits at their resonant frequencies on the two transmission lines between the hybrid rings. Each pair of cavities is tuned to the same frequency so that the short circuits produced by the aural notch cavities reflect the aural frequency power which add in phase at the antenna terminal of the hybrid ring. Likewise, the power in the lower side-band frequencies is reflected back into the dummy load. Of course, the 4.5-mc upper side bands of the video transmitter are reflected back into the dummy load and are not transmitted to the antenna.

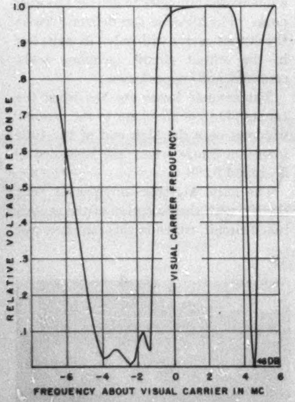


Fig. 14. Typical frequency response curve of combined diplexer and lower side-band filter. Data taken at 730 mc

The frequency response of the combined unit through the visual input to the antenna terminal is shown in Fig. 14. The side-band filter at uhf is not proportionally smaller than the equivalent filter at very high frequency, as it requires the same physical volume to achieve the lower side-band attenuation shown at both uhf and very high frequency. The sound rejection filters are made of copper-clad invar wave guide to reduce temperature drift to a minimum.

The power output of the transmitter varies with the beam voltage, and Fig. 15 shows the narrow-band power output and broad-band power output with a typical tube. In the theoretical investigation, it was shown that the power output and efficiency should be the same for broad-band or narrow-band operation. However, in broad-band operation of the center gap, the power gain would drop so that more radio frequency voltage at the input gap would be required for maximum bunching at the output gap for the same power output. This power gain did drop and the efficiency was less as shown on the data when 300 watts were dissipated in the center circuit's dummy load. However, data on the next transmitter installed have shown much less difference between the two methods of operation and with less power dissipated in the center-tuned circuits for the same bandwith.

The linearity of the visual system is shown in Fig. 16. The modulated amplifier contributes all of the white compression at low-power levels, and the klystron contributes the blacker-than-black compression at 8 kw. Less drive will reduce both the compression and the power output. In normal station operation, a stabilizing amplifier is used to set the synchronizing signal amplitude as a percentage of peak carrier.

The sound transmitter exciter measured about −60-db frequency-modulation noise and, adding the klystron amplifier, it was still −60 db. Amplitude-modulation noise on the sound carrier measured approximately −50 db. The video transmitter and monitor receiver have a rise time of 0.12 microsecond with a 100-kc square wave.

References

1. Klystron Amplificatrur de 5 KW à large bande passante, P. Guénard, B. Epsztein, P. Cahour. *Annales de Radioélectricité*, Compagnies Francaises Associées de T.S.F., Paris, France, vol. VI, 1951, p. 24.

2. A Simple Crystal Discriminator for FM Oscillator Stabilization, John Ruston. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 39, no. 9, July 1951.

3. Report No. 12.2.15/215, W. Chesnut. Allen B. Du Mont Laboratories, Inc., Clifton, N. J.]

4. Method of Measuring the Properties of Ionized Gases at High Frequencies, C. Sanborn, C. Brown, David J. Rose. *Report No. 222*, Massachusetts Institute of Technology, Cambridge, Mass.

5. Report 12.2.20/25, Allen B. Du Mont Laboratories, Inc., Clifton, N. J.

◆

No Discussion

Frequency Response of a Resonant Dielectric Amplifier

G. W. PENNEY
FELLOW AIEE

E. A. SACK
STUDENT MEMBER AIEE

E. R. WINGROVE
NONMEMBER AIEE

Synopsis: A fundamental limitation on the steady-state signal frequency response of a series-resonant dielectric voltage amplifier is the time required for the resonant circuit to adjust to a change in tuning. This paper analyzes this limitation for the case where nonlinear resonance effects may be neglected. Typically, this limitation may require that the power or carrier frequency be several hundred times the highest signal frequency, if the gain of the amplifier is to be constant.

THE dielectric voltage amplifier shown in Fig. 1 utilizes a nonlinear dielectric (ferroelectric) capacitor as part of a series resonant circuit.[1] This circuit is excited by a carrier source whose voltage and frequency are constant. The capacitance of the nonlinear capacitor is varied by the signal voltage to be amplified. Of course, a change in the capacitance alters the tuning of the resonant circuit, and since the carrier frequency is constant, the circulating carrier current in the resonant circuit is modulated by the signal. The voltages appearing across the elements of the circuit will likewise be modulated. If one of these voltages, usually that across the inductance, is demodulated, a voltage output is obtained which, in a properly designed amplifier, has similar waveshape but larger amplitude than the input signal. Thus a voltage gain is realized.

The average stored energy in a high-Q circuit operating near resonance will be many times the energy which can be added by the source or dissipated in circuit losses during a cycle of the circulating current. There is a definite average stored energy associated with every value of tuning in the region of resonance. If a change in tuning requiring a relatively large change in average stored energy occurs during a few cycles of the circulating current, it is impossible for the source to supply or the circuit losses to dissipate the required change in energy during this time. Since this energy is related to the circulating current amplitude, it is impossible for the current to adjust to the new steady-state value required by the change in tuning during the time in which the change occurs.

If the signal frequency, and thus the frequency of the tuning variation, is low, the circulating current may very nearly approach the steady-state amplitude required by the tuning at each instant. However, as the signal frequency approaches some fraction of the carrier frequency, the circulating current will no longer be able to follow the variations in tuning and the carrier-current modulation must decrease in amplitude and fall behind the signal. Since the signal output of the amplifier is proportional to the carrier current modulation, this output must also decrease in amplitude and lag in phase for higher signal frequencies.

In addition to the fundamental limitation considered in this paper, the frequency response of the amplifier of Fig. 1 may also be restricted by the signal input and demodulator circuits. It is generally possible, however, to design these circuits so that the drop off in the circulating-current modulation is the controlling limitation on the high-frequency response.

In order to obtain quantitative information on the relation between the carrier modulation amplitude and the signal frequency, an investigation is made into the response of a linear-element resonant circuit having sinusoidally varying capacitance. Of course, in the dielectric amplifier, the variable capacitance is actually nonlinear. However, this paper analyzes the case where the carrier-frequency voltage drop across the nonlinear capacitor is sufficiently small that it causes negligible variation in the total circuit capacitance. The results of the linear analysis should be of some value, even when the nonlinearity in the resonant loop is no longer negligible, since the average stored energy for the resonant circuit must still change as the circuit tuning varies.

Nomenclature

V = peak value of the applied carrier potential
i = instantaneous value of the circulating current
q = instantaneous value of the charge on the resonant circuit capacity
$q_0, q_1, q_2, \ldots$ = coefficients of the power-series expansion for q
ω_s = signal angular frequency
ω_c = carrier angular frequency
ω_N = natural angular frequency of the resonant circuit for a circuit capacitance C
t = time
ϕ_s = phase angle between the circulation-current modulation envelope and the circuit-capacitance modulation
ϕ_c = phase angle between the circulating current and the applied carrier potential
R = resistance in the resonant circuit
L = inductance in the resonant circuit
$c(t)$ = instantaneous value of resonant circuit capacitance
C = average capacitance in the resonant circuit
Q = quality factor of the resonant circuit
m = modulation factor for the resonant circuit capacitance
M = modulation factor of the circulating current
M_0 = modulation factor for very low signal frequencies
x, y = variables of the general differential equation
a, b, d = coefficients of the general differential equation
u, v = linear functions of x
f_{k-1} = right hand side of the $(k-1)$st equation in the array
$\beta, \gamma, \delta, \alpha_1, \beta_1, \gamma_1, \delta_1, D_1, E_1, F_1, D_2, E_2, F_2, D_3, E_3, F_3, D_4, \ldots$ = constants used to avoid rewriting unwieldy expressions

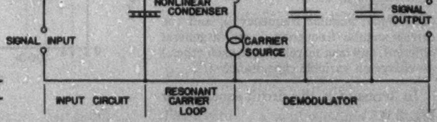


Fig. 1. A resonant-circuit dielectric amplifier

Paper 54-160, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 19, 1953; made available for printing December 14, 1953.

G. W. PENNEY, E. A. SACK, and E. R. WINGROVE are with the Carnegie Institute of Technology, Pittsburgh, Pa.

This work has been done in part under Office of Naval Research Contract *N9-onr-30308*, and in part by E. A. Sack and E. R. Wingrove in partial fulfillment of the requirements for the degree of Doctor in Philosophy at the Carnegie Institute of Technology.

The authors wish to acknowledge the advice and suggestions of Dr. W. A. Altar and Dr. C. W. Helstrom of the Westinghouse Electric Corporation. Dr. Helstrom suggested the general method of solution.

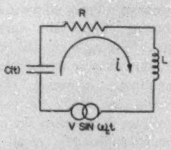


Fig. 2. Resonant loop

Method of Solution

The differential equation for the circuit under investigation is second order, linear, with a time-varying parameter. This equation cannot be solved by ordinary methods. The solution presented here utilizes an assumed infinite series in powers of the modulation factor m where the coefficients of the series are functions of time to be determined. Application of this series transforms the task of solving the nonconstant-parameter equation into one of solving an array of constant-parameter equations, one at a time. In the array, the solution of the kth equation yields the kth coefficient of the power series, and makes possible the solution of the next member of the array. In the case of the dielectric amplifier, the power series converges rapidly, so that it is necessary to solve only a few of the equations.

The method of solution is not restricted to the case investigated here, but can also be applied to a second-order circuit where the variation in either inductance, resistance, or capacitance can be expressed in a rapidly convergent power series of the form

$$L(t)=L(1+m\sin\omega_s t+m^2\sin^2\omega_s t+\ldots)$$
$$R(t)=R(1+m\sin\omega_s t+m^2\sin^2\omega_s t+\ldots)$$
$$1/C(t)=(1/C)(1+m\sin\omega_s t+m^2\sin^2\omega_s t+\ldots)$$

The electric driving force in the circuit may be other than sinusoidal, although more complicated forms may greatly increase the work involved in evaluating the coefficients of the power series.

Differential Equation and Solution

In the circuit of Fig. 2

L = constant inductance
R = constant resistance
$c(t)$ = capacitance which varies as
$c(t)=C(1+m\sin\omega_s t)\quad m<1$ (1)
$V\sin\omega_c t$ = electric driving force or "carrier supply"
(The signal angular frequency ω_s and the carrier angular frequency ω_c are in general different, and bear no relation to each other.)
q = charge on variable capacitance $c(t)$

In terms of q, the circuit equation for Fig. 2 is

$$L\ddot{q}+R\dot{q}+\frac{1}{C(1+m\sin\omega_s t)}q=V\sin\omega_c t \tag{2}$$

or, multiplying equation 2 by

$(1+m\sin\omega_s t)/L$

$$(1+m\sin\omega_s t)\ddot{q}+(1+m\sin\omega_s t)\frac{R}{L}\dot{q}+\frac{1}{LC}q=(1+m\sin\omega_s t)\frac{V}{L}\sin\omega_c t \tag{3}$$

The assumed solution to equation 3 is an infinite series of increasing powers of the capacitance modulation factor m

$$q=q_0+q_1m+q_2m^2+q_3m^3+q_4m^4+\ldots \tag{4}$$

where $q_0, q_1, q_2, \ldots$ are functions of time to be determined.

Substituting equation 4 into 3, and equating the coefficients of like powers of m, yields the infinite array of equations

$$\ddot{q}_0+\frac{R}{L}\dot{q}_0+\frac{1}{LC}q_0=\frac{V}{L}\sin\omega_c t \tag{5.0}$$

$$\ddot{q}_1+\frac{R}{L}\dot{q}_1+\frac{1}{LC}q_1=\left(\frac{V}{L}\sin\omega_c t-\ddot{q}_0-\frac{R}{L}\dot{q}_0\right)\sin\omega_s t \tag{5.1}$$

$$\ddot{q}_2+\frac{R}{L}\dot{q}_2+\frac{1}{LC}q_2=-\left(\ddot{q}_1+\frac{R}{L}\dot{q}_1\right)\sin\omega_s t \tag{5.2}$$

$$\vdots$$

$$\ddot{q}_k+\frac{R}{L}\dot{q}_k+\frac{1}{LC}q_k=-\left(\ddot{q}_{k-1}+\frac{R}{L}\dot{q}_{k-1}\right)\sin\omega_s t \tag{5.k}$$

$$\vdots$$

The right-hand side of each of the equations, except equation 5.0, may be written in a different form, by using an equality obtained in each case from the preceding equation. Multiplying equation 5.0 by ($\sin\omega_s t$) and rearranging terms yields an expression for the right-hand side of equation 5.1 which involves only q_0 and not its derivatives

$$\left(\frac{V}{L}\sin\omega_c t-\ddot{q}_0-\frac{R}{L}\dot{q}_0\right)\sin\omega_s t=\frac{1}{LC}q_0\sin\omega_s t \tag{6}$$

Proceeding in this manner for the entire array gives

$$\ddot{q}_0+\frac{R}{L}\dot{q}_0+\frac{1}{LC}q_0=\frac{V}{L}\sin\omega_c t \tag{7.0}$$

$$\ddot{q}_1+\frac{R}{L}\dot{q}_1+\frac{1}{LC}q_1=\frac{1}{LC}q_0\sin\omega_s t \tag{7.1}$$

$$\ddot{q}_2+\frac{R}{L}\dot{q}_2+\frac{1}{LC}q_2=\left(\frac{1}{LC}q_1-\frac{1}{LC}q_0\sin\omega_c t\right)\sin\omega_s t \tag{7.2}$$

$$\vdots$$

$$\ddot{q}_k+\frac{R}{L}\dot{q}_k+\frac{1}{LC}q_k=\left(\frac{1}{LC}q_{k-1}-f_{k-1}\right)\sin\omega_s t \tag{7.k}$$

$$\vdots$$

where (f_{k-1}) is the right-hand side of the ($k-1$) st equation. Each of the equations in the array is linear and has constant coefficients. Equation 7.0 may be solved for q_0 in terms of the parameters L, R, C, V, ω_c and t. This solution for q_0 can be substituted into equation 7.1, which can then be solved for q_1 in terms of L, R, C, V, ω_c, ω_s, and t. In general, the solution to the ($k-1$)st equation leads to the solution of the kth equation in the same way.

The right-hand side of each equation in the array can be expressed as a sum of a finite number of sinusoidal components.

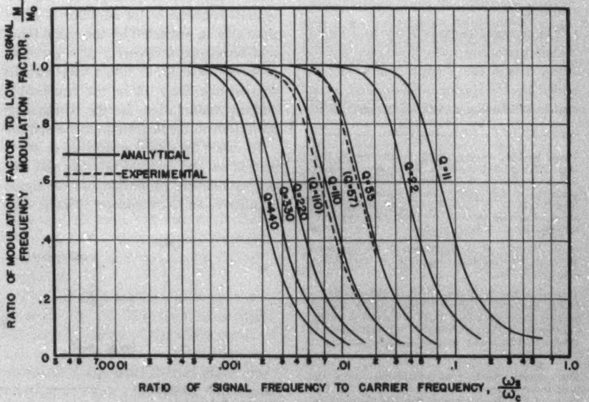


Fig. 3. Modulation amplitude characteristics

The solutions of the equations may be obtained by either standard differential-equation or a-c theory methods. Since each equation is linear, its solution is the sum of the solutions for each of the right-hand-side driving components

Solving equation 7.0 for q_0 gives

$$q_0=\frac{\frac{V}{R\omega_c}}{1+\left(\frac{1}{\omega_c RC}-\frac{\omega_c L}{R}\right)^2}\left\{-\cos\omega_c t+\left(\frac{1}{\omega_c RC}-\frac{\omega_c L}{R}\right)\sin\omega_c t\right\} \tag{8}$$

Inserting equation 8 into the right-hand side of equation 7.1, and solving for q_1 yields

$$q_1=D_1\sin(\omega_c+\omega_s)t+D_2\cos(\omega_c+\omega_s)t+ D_3\sin(\omega_c-\omega_s)t+D_4\cos(\omega_c-\omega_s)t \tag{9}$$

where the coefficients are defined by equation 27 in the Appendix. Whereas, the solution to q_0 contributes only carrier-frequency terms, q_1, adds first side-band components.

If now equation 9 is substituted into the right-hand side of equation, 7.2 the equation can be solved for q_2, and the solution has the form

$$q_2=E_1\sin\omega_c t+E_2\cos\omega_c t+E_3\sin(\omega_c+2\omega_s)t+E_4\cos(\omega_c+2\omega_s)t+ E_5\sin(\omega_c-2\omega_s)t+E_6\cos(\omega_c-2\omega_s)t \tag{10}$$

where E_1 through E_6 are constants. Thus, equation 7.2 contributes more carrier-frequency components, as well as the secondary side-band terms. It can be shown that the solution to equation 7.3 will include additional first side-band components, as well as tertiary side bands. From this, a practical limitation of this form of solution becomes apparent. If m is not sufficiently small to make these additional carrier and primary side-band terms negligible, it becomes necessary to evaluate more than the first two or three coefficients of the power series. Although the procedure of solution for equations 7.3 7.4, etc., remains the same, the number of terms on the right hand side of each differential equation becomes inconveniently large.

In the case of several practical dielectric amplifiers m is equal to or less than 0.006. Thus, a solution which includes only the first two terms of the power series is sufficiently accurate for the purpose of this investigation.

$$q\cong q_0+mq_1 \tag{11}$$

The circulating current in the resonant circuit is

$$i=\frac{dq}{dt}\cong\frac{dq_0}{dt}+m\frac{dq_1}{dt} \tag{12}$$

From equations 8 and 9 the form of the expression for the current is

$$i=\frac{\frac{V}{R}}{1+\left(\frac{1}{\omega_c RC}-\frac{\omega_c L}{R}\right)^2}\left\{\sin\omega_c t+F_1\cos\omega_c t+\frac{m}{2}\frac{1}{\omega_c RC}\left[F_2\sin(\omega_c+\omega_s)t+F_3\cos(\omega_c+\omega_s)t+F_4\sin(\omega_c-\omega_s)t+F_5\cos(\omega_c-\omega_s)t\right]\right\} \tag{13}$$

where the coefficients F_1 through F_5 are listed in the Appendix. Equation 13 is not a convenient form from which to determine the amplitude of the modulation envelope. As is shown in the Appendix, by again neglecting terms multiplied by powers of m higher than the first, it is possible to rewrite equation 13 as

$$i=\frac{\frac{V}{R}}{\sqrt{1+\left(\frac{1}{\omega_c RC}-\frac{\omega_c L}{R}\right)^2}}\times\left\{1+\frac{m}{2}\frac{1}{\omega_c RC}\frac{\alpha_1\delta_1+\beta_1\gamma_1}{\sqrt{(\beta_1^2+\delta_1^2)(\alpha_1^2+\gamma_1^2)}}\times\sin(\omega_s t+\phi_s)\right\}\sin(\omega_c t+\phi_c) \tag{14}$$

where

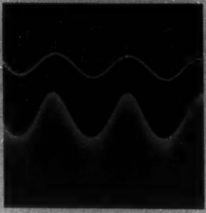


Fig. 5. Top: Input signal. Bottom: Modulated carrier

$$\alpha_1=\left(1+\frac{\omega_s}{\omega_c}\right)$$

$$\beta_1=\left(1-\frac{\omega_s}{\omega_c}\right)$$

$$\gamma_1=\left[\frac{1}{\omega_c RC}-\left(1+\frac{\omega_s}{\omega_c}\right)^2\frac{\omega_c L}{R}\right]$$

$$\delta_1=\left[\frac{1}{\omega_c RC}-\left(1-\frac{\omega_s}{\omega_c}\right)^2\frac{\omega_c L}{R}\right]$$

and ϕ_c is the phase angle of the carrier. The angle by which the circulating current modulation envelope lags the capacitance variation is

$$\phi_s=\tan^{-1}\frac{\beta_1\gamma_1-\alpha_1\delta_1}{\beta_1\alpha_1+\gamma_1\delta_1} \tag{15}$$

In terms of the circuit natural frequency $\omega_N=1/\sqrt{LC}$ and quality factor $Q=(\omega_c L)/R$, equation 14 can also be written

$$i=\frac{\frac{V}{R}}{\sqrt{1+Q^2\left(\frac{\omega_N^2}{\omega_c^2}-1\right)^2}}\times\left\{1+\frac{m}{2}Q\frac{\omega_N^2}{\omega_c^2}\frac{\alpha\delta+\beta\gamma}{\sqrt{(\beta^2+\delta^2)(\alpha^2+\gamma^2)}}\times\sin(\omega_s t+\phi_s)\right\}\sin(\omega_c t+\phi_c) \tag{16}$$

where now

$$\alpha=\left(1+\frac{\omega_s}{\omega_c}\right)$$

$$\beta=\left(1-\frac{\omega_s}{\omega_c}\right)$$

$$\gamma=Q\left[\frac{\omega_N^2}{\omega_c^2}-\left(1+\frac{\omega_s}{\omega_c}\right)^2\right]$$

$$\delta=Q\left[\frac{\omega_N^2}{\omega_c^2}-\left(1-\frac{\omega_s}{\omega_c}\right)^2\right]$$

Modulation Factor

The quantity of interest in equation 16 is the amplitude of the modulation envelope M

$$M=\frac{m}{2}Q\frac{\omega_N^2}{\omega_c^2}\frac{\alpha\delta+\beta\gamma}{\sqrt{(\beta^2+\delta^2)(\alpha^2+\gamma^2)}} \tag{17}$$

The limiting value of equation 17, as the

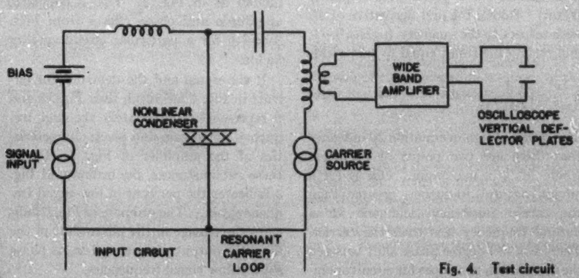


Fig. 4. Test circuit

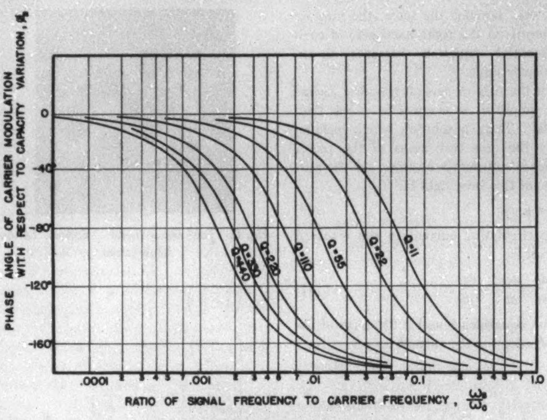


Fig. 6. Modulation phase characteristics

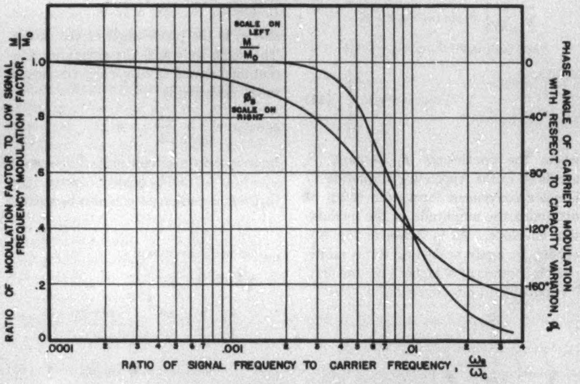


Fig. 7. Comparison of modulation amplitude and phase characteristic

signal frequency becomes very small in comparison with the carrier frequency, is

$$M_0 = m\frac{\omega_N^2}{\omega_c^2}\,\frac{Q^2\left(\dfrac{\omega_N^2}{\omega_c^2}-1\right)}{1+Q^2\left(\dfrac{\omega_N^2}{\omega_c^2}-1\right)^2} \tag{18}$$

Of primary importance is the ratio

$$\frac{M}{M_0} = \frac{1+Q^2\left(\dfrac{\omega_N^2}{\omega_c^2}-1\right)^2}{2Q\left(\dfrac{\omega_N^2}{\omega_c^2}-1\right)}\,\frac{\alpha\delta+\beta\gamma}{\sqrt{(\beta^2+\delta^2)(\alpha^2+\gamma^2)}} \tag{19}$$

Before plotting M/M_0 versus ω_s/ω_c for various values of the circuit quality factor $(\omega_c L)/R$, it is necessary to pick a value for $[(\omega_N^2/\omega_c^2)-1]$. This quantity determines where the point of operation for the circuit lies on its resonance curve. For the dielectric amplifier, $(\omega_N^2/\omega_c^2)-1$ is generally chosen to maximize gain and should be picked, therefore, so that M_0 is a maximum for a given circuit quality factor. Taking the first derivative of M_0 with respect to the quantity $(\omega_N^2/\omega_c^2)-1$, and setting the result equal to zero yields

$$\left(\frac{\omega_N^2}{\omega_c^2}-1\right)_{\max' M_0} = \pm\frac{1}{Q} \tag{20}$$

The choice of sign in equation 20 indicates that there are two points of operation where M_0 is a maximum. One is for a circuit natural frequency greater than the carrier frequency, and one for a natural frequency less than the carrier. There is a 180-degree phase shift between the modulation envelopes for circuit operation at each of these two points, but M/M_0 remains the same for either sign.

For $(\omega_N^2/\omega_c^2)-1=\pm 1/Q$, equation 19 becomes

$$\frac{M}{M_0} = \left|\frac{\alpha\delta+\beta\gamma}{\sqrt{(\beta^2+\delta^2)(\alpha^2+\gamma^2)}}\right| \tag{21}$$

Response and Experimental Checks

Fig. 3 shows the drop off in modulation amplitude with increasing signal frequency for several circuit quality factors. The solid curves were obtained from equation 21 while the dotted curves were taken experimentally, using the circuit shown in Fig. 4.

The test circuit used is shown in Fig. 4. This circuit is essentially one form of the resonant dielectric amplifier, without the output demodulator. The input circuit, consisting of a signal generator and a carrier blocking choke, places a constant-amplitude signal across the nonlinear dielectric capacitor for all signal frequencies in the test range. The signal amplitude is limited to a value which results in the desired variation in total circuit capacitance, $C(1+m\sin\omega_s t)$. The carrier potential is sufficiently small that the circuit capacitance remains essentially constant throughout a cycle of the carrier. The modulation envelope of the voltage drop across L_T is essentially the same as the modulation envelope of the circulating current, so long as M is small. The voltage across a link loosely coupled to L_T is amplified and applied directly to the vertical deflection plates of an oscilloscope. The ratio M/M_0 is computed from measurements of the envelope as it appears on the screen of the oscilloscope. Fig. 5 shows this envelope and the modulating signal.

Fig. 6 shows the variation in the phase angle between the circulating current modulation envelope and the circuit-capacitance modulation, as the ratio of signal-to-carrier frequency changes. Curves are shown for the same quality factors as in Fig. 3. Fig. 7 compares amplitude and phase curves from Figs. 3 and 6 for a particular circuit quality factor.

If the signal and the demodulator circuits in Fig. 1 are ideal, then Figs. 3 and 6 represent approximately the high-frequency-signal gain and phase characteristics of the amplifier of Fig. 1. Under these circumstances, the ordinate of Fig. 3 indicates the per cent of low-signal frequency gain. The ordinate of Fig. 6 indicates the change in the phase shift of the amplifier output with respect to its phase shift at low signal frequencies.

In this analysis only two terms of the assumed power series have been used in finding M_0. In order to check the accuracy of this approximation, the low-signal frequency gain of a dielectric amplifier as calculated from M_0 can be compared with a more direct analysis of this gain from a previous paper.[1] The latter also treats the resonant loop as linear, but is an exact analysis of the case where the signal frequency approaches zero. For a given stage, the gain as calculated from the first two terms of the assumed series agrees with the result from the direct analysis to within 2 per cent, for a selection of circuit quality factors in the range 22 to 550.

Conclusions

The solution shows that the fundamental limitation on the signal frequency response of a resonant-circuit dielectric amplifier is the time required for the resonant loop to adjust to a change in tuning. If this is the only limitation, then for an amplifier with a carrier-circuit quality factor of 110 to have a flat response to 10,000 cycles per second, a carrier frequency on the order of 7 megacycles must be employed. It can be seen from Fig. 4 that the signal frequency at the high-frequency half-power-output point is approximately 0.7 of the ratio of the carrier frequency to the circuit quality factor. If an amplifier is to have negligible phase shift at a particular signal frequency, it must utilize an even higher carrier frequency than that required for a flat amplitude response.

Appendix

Solution for the Coefficients of the Assumed Power Series

From equations 1 through 7 the substitution of the power series

$$q=q_0+q_1m+q_2m^2+q_3m^3+q_4m^4+\ldots \qquad (4)$$

leads to the array

$$\ddot{q}_0+\frac{R}{L}\dot{q}_1+\frac{1}{LC}q_0=\frac{V}{L}\sin\omega_c t \qquad (7.0)$$

$$\ddot{q}_1+\frac{R}{L}\dot{q}_1+\frac{1}{LC}q_1=\frac{1}{LC}q_0\sin\omega_s t \qquad (7.1)$$

$$\ddot{q}_2+\frac{R}{L}\dot{q}_2+\frac{1}{LC}q_2=\left(\frac{1}{LC}q_1-\frac{1}{LC}q_0\sin\omega_s t\right)\times \sin\omega_s t \qquad (7.2)$$

$$\vdots$$

$$\ddot{q}_k+\frac{R}{L}\dot{q}_k+\frac{1}{LC}q_k=\left(\frac{1}{LC}q_{k-1}-f_{k-1}\right)\times \sin\omega_s t \qquad (7.k)$$

$$\vdots$$

The right-hand side of each member of the array can be expressed as a sum of sine or cosine terms having frequencies which are sums or differences of multiples of ω_c and ω_s. The general form of any member of the array set equal to only one of its right-hand side sinusoids is either

$$\frac{d^2y}{dx^2}+a\frac{dy}{dx}+by=d\sin(u+v) \qquad (22)$$

or

$$\frac{d^2y}{dx^2}+a\frac{dy}{dx}+by=d\cos(u+v) \qquad (23)$$

where u and v are linear functions of x. The general form of the steady-state solution to equation 22 is

$$y=\frac{d}{[b-(u'+v')^2]^2+a^2(u'+v')^2}\times \left\{[b-(u'+v')^2]\sin(u+v)- a(u'+v')\cos(u+v)\right\} \qquad (24)$$

while the general steady state solution to equation 23 is

$$y=\frac{d}{[b-(u'+v')^2]^2+a^2(u'+v')^2}\times \left\{a(u'+v')\sin(u+v)+ [b-(u'+v')^2]\cos(u+v)\right\} \qquad (25)$$

Equation 7.0 has the general form shown by equation 22. Hence, from equation 24, the solution to equation 7.0 is

$$q_0=\frac{\dfrac{V}{\omega_c R}}{1+\left(\dfrac{1}{\omega_c RC}-\dfrac{\omega_c L}{R}\right)^2} \left\{-\cos\omega_c t+\left(\frac{1}{\omega_c RC}-\frac{\omega_c L}{R}\right)\sin\omega_c t\right\} \qquad (8)$$

Substituting equation 8 into equation 7.1 and rewriting the right-hand-side terms as functions of $(\omega_c+\omega_s)$ and $(\omega_c-\omega_s)$ yields

$$\ddot{q}_1+\frac{R}{L}\dot{q}_1+\frac{1}{LC}q_1=\left[\frac{\dfrac{V}{2L^2C}}{\left(\dfrac{1}{LC}-\omega_c^2\right)^2+\dfrac{R^2}{L^2}\omega_c^2}\right] \left[-\frac{R}{L}\omega_c\sin(\omega_c+\omega_s)t-\left(\frac{1}{LC}-\omega_c^2\right)\times \cos(\omega_c+\omega_s)t+\frac{R}{L}\omega_c\sin(\omega_c-\omega_s)t+ \left(\frac{1}{LC}-\omega_c^2\right)\cos(\omega_c-\omega_s)t\right] \qquad (26)$$

Since the principle of superposition applies, the steady-state solution to equation 26 is the sum of the steady state solutions to the four equations formed by setting

$$\ddot{q}_1+\frac{R}{L}\dot{q}_1+\frac{1}{LC}q_1$$

equal to each of the right hand side terms of equation 26. Each of these four equations has either the general form of equation 22 or 23. Hence, the solution to each equation can be obtained from either equation 24 or 25. Forming these solutions and collecting terms

$$q_1=\frac{\dfrac{V}{2\omega_c R}}{1+F_1^2}\frac{1}{\omega_c RC}\left[\frac{F_1\gamma_1-\alpha_1}{\gamma_1^2+\alpha_1^2}\sin(\omega_c+\omega_s)t+ \frac{-\gamma_1-\alpha_1F_1}{\gamma_1^2+\alpha_1^2}\cos(\omega_c+\omega_s)t+ \frac{-F_1\delta_1+\beta_1}{\delta_1^2+\beta_1^2}\sin(\omega_c-\omega_s)t+ \frac{\delta_1+\beta_1F_1}{\delta_1^2+\beta_1^2}\cos(\omega_c-\omega_s)t\right] \qquad (27)$$

where

$$\alpha_1=\left(1+\frac{\omega_s}{\omega_c}\right);\quad \beta_1=\left(1-\frac{\omega_s}{\omega_c}\right)$$

$$\gamma_1=\left[\frac{1}{\omega_c RC}-\left(1+\frac{\omega_s}{\omega_c}\right)^2\frac{\omega_c L}{R}\right]$$

$$\delta_1=\left[\frac{1}{\omega_c RC}-\left(1-\frac{\omega_s}{\omega_c}\right)^2\frac{\omega_c L}{R}\right]$$

$$F_1=\left(\frac{1}{\omega_c RC}-\frac{\omega_c L}{R}\right)$$

Forming

$$q\cong q_0+mq_1 \qquad (11)$$

and taking the derivative

$$i=\frac{dq}{dt}\cong\frac{dq_0}{dt}+m\frac{dq_1}{dt} \qquad (12)$$

From equations 8 and 27

$$i=\frac{\dfrac{V}{R}}{1+F_1^2}\left\{\sin\omega_c t+F_1\cos\omega_c t+ \frac{m}{2}\frac{1}{\omega_c RC}[F_2\sin(\omega_c+\omega_s)t+F_3\times \cos(\omega_c+\omega_s)t+F_4\sin(\omega_c-\omega_s)t+ F_5\cos(\omega_c-\omega_s)t]\right\} \qquad (13)$$

where

$$F_1=\left(\frac{1}{\omega_c RC}-\frac{\omega_c L}{R}\right)$$

$$F_2=\frac{\alpha_1F_1\gamma_1-\alpha_1^2}{\gamma_1^2+\alpha_1^2}$$

$$F_3=\frac{-\alpha_1\gamma_1-\alpha_1^2F_1}{\gamma_1^2+\alpha_1^2}$$

$$F_4=\frac{-\beta_1F_1\delta_1+\beta_1^2}{\delta_1^2+\beta_1^2}$$

$$F_5=\frac{\beta_1\delta_1+\beta_1^2F_1}{\delta_1^2+\beta_1^2}$$

Simplification of Equation 13

The form of equation 13 is

$$i=\frac{\dfrac{V}{R}}{1+F_1^2}\left\{\sin\omega_c t+F_1\cos\omega_c t+\frac{m}{2}\frac{1}{\omega_c RC}\times [F_2\sin(\omega_c+\omega_s)t+F_3\cos(\omega_c+\omega_s)t+ F_4\sin(\omega_c-\omega_s)t+F_5\cos(\omega_c-\omega_s)t]\right\} \qquad (13)$$

Expanding and collecting terms

$$i=\frac{\frac{V}{R}}{1+F_1^2}\left\{\left[1+\frac{m}{2}\frac{1}{\omega_cRC}(F_5-F_3)\sin\omega_st+\right.\right.$$
$$\left.\frac{m}{2}\frac{1}{\omega_cRC}(F_2+F_4)\cos\omega_st\right]\times$$
$$\sin\omega_ct+\left[F_1+\frac{m}{2}\frac{1}{\omega_cRC}(F_2-F_4)\times\right.$$
$$\left.\sin\omega_st+\frac{m}{2}\frac{1}{\omega_cRC}(F_3+F_5)\cos\omega_st\right]\times$$
$$\left.\cos\omega_ct\right\} \quad (28)$$

Combining the sin ω_ct and cos ω_ct terms

$$i=\frac{\frac{V}{R}}{1+F_1^2}\left\{(1+F_1^2)+\frac{m}{\omega_cRC}[(F_5-F_3)+\right.$$
$$F_1(F_2-F_4)]\sin\omega_st+\frac{m}{\omega_cRC}\times$$
$$[(F_2+F_4)+F_1(F_3+F_5)]\cos\omega_st+$$
$$\frac{m^2}{4(\omega_cRC)^2}[(F_2-F_4)^2+(F_5-F_3)^2]\times$$
$$\sin^2\omega_st+\frac{m^2}{4(\omega_cRC)^2}[(F_2+F_4)^2+(F_3+$$
$$F_5)^2]\cos^2\omega_st+\frac{m^2}{2(\omega_cRC)^2}\times$$
$$[(F_5-F_3)(F_2+F_4)+(F_2-$$
$$\left.F_4)(F_3+F_5)]\sin\omega_st\cos\omega_st\right\}^{\frac{1}{2}}\times$$
$$\sin(\omega_ct+\phi_c) \quad (29)$$

Neglecting terms multiplied by m^2 and combining sin ω_st and cos ω_st

$$i=\frac{\frac{V}{R}}{\sqrt{1+F_1^2}}\left[1+\frac{m}{1+F_1^2}\frac{1}{\omega_cRC}\right.$$
$$\sqrt{[(F_2+F_4)+F_1(F_3+F_5)]^2+}$$
$$[(F_5-F_3)+F_1(F_2-F_4)]^2\times$$
$$\left.\sin(\omega_st+\phi_s)\right]^{\frac{1}{2}}\sin(\omega_ct+\phi_c) \quad (30)$$

where

$$\phi_s=\tan^{-1}\frac{(F_2+F_4)+F_1(F_3+F_5)}{(F_5-F_3)+F_1(F_2-F_4)} \quad (31)$$

Writing

$$\left[1+\frac{m}{1+F_1^2}\frac{1}{\omega_cRC}\times\right.$$
$$\sqrt{[(F_2+F_4)+F_1(F_3+F_5)]^2+}$$
$$[(F_5-F_3)+F_1(F_2-F_4)]^2\times$$
$$\left.\sin(\omega_st+\phi_s)\right]^{\frac{1}{2}}$$

in a binomial expansion and again neglecting all terms which are multiplied by powers of m higher than the first

$$i=\frac{\frac{V}{R}}{\sqrt{1+F_1^2}}\left\{1+\frac{m}{2(1+F_1^2)}\frac{1}{\omega_cRC}\times\right.$$
$$\sqrt{[(F_2+F_4)+F_1(F_3+F_5)]^2+}$$
$$[(F_5-F_3)+F_1(F_2-F_4)]^2\times$$
$$\left.\sin(\omega_st+\phi_s)\right\}\sin(\omega_ct+\phi_c) \quad (32)$$

Substituting the values given for F_1, F_2, F_3, F_4, and F_5

$$i=\frac{\frac{V}{R}}{\sqrt{1+\left(\frac{1}{\omega_cRC}-\frac{\omega_cL}{R}\right)^2}}\times$$
$$\left[1+\frac{m}{2}\frac{1}{\omega_cRC}\frac{\beta_1\gamma_1+\alpha_1\delta_1}{\sqrt{(\gamma_1^2+\alpha_1^2)(\delta_1^2+\beta_1^2)}}\times\right.$$
$$\left.\sin(\omega_st+\phi_s)\right]\sin(\omega_ct+\phi_c \quad (14)$$

$$\phi_s=\tan^{-1}\frac{\beta_1\gamma_1-\alpha_1\delta_1}{\alpha_1\beta_1+\gamma_1\delta_1} \quad (15)$$

Reference

1. DIELECTRIC AMPLIFIERS, G. W. Penney, J. R. Horsch, E. A. Sack. *AIEE Transactions*, vol. 72, pt. I, March 1953, pp. 68–79.

No Discussion

Vol
Pt.
195

An Automatic Voltage Control System for Electrical Precipitators

H. J. HALL
NONMEMBER AIEE

A FUNDAMENTALLY new system of precipitator automatic voltage control has been developed for use in the Cottrell process of electrically removing suspended matter from gases. The new system is based on extensive field measurements with respect to the actual operating requirements necessary to obtain fundamentally sound performance. This, coupled with flexibility in design, ensures that optimum electrical conditions for maximum collection efficiency will be maintained automatically and continuously for a wide range of precipitator conditions.

The need for automatic voltage control in electrical precipitation equipment has long been appreciated. Proposed systems have been discussed in the literature[1,2,3] from time to time, but none of these has gained acceptance or been used to any extent. The operation of earlier systems depended on one or more of the following:

1. Onset of precipitator sparking; i.e., the regulator takes hold at the first spark and essentially maintains the precipitator operating voltage at, or just below, the sparking potential.
2. Predetermined values of precipitator voltage or current.
3. A line input voltage which is kept constant.

The complexity of the factors which determine precipitator operating voltages and currents and the fact that they may change frequently, and sometimes over wide ranges, makes it highly impractical to use either the voltage or the current as a primary control factor in an automatic regulating system. In addition, the control elements in previous systems have often been large, expensive, and rather complicated. Consequently, such systems have not been incorporated generally into electrical precipitation equipment in this country.

On the other hand, the value of a properly functioning automatic regulator in continuously maintaining maximum precipitator performance is reflected in increased long-term collection efficiency and reduced operating expense. The current public and engineering emphasis on the reduction of atmospheric pollution further stresses the importance of automatically maintaining maximum electrical energization in precipitators subject to variable load conditions.

Basic Design Requirements

As a necessary and important preliminary step in developing a suitable automatic voltage control system, field measurements of actual operating characteristics and investigations of the requirements for best efficiency were carried out on a number of industrial precipitators. These

---

Paper 54-175, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 21, 1953; made available for printing November 27, 1953.

H. J. HALL is with the Research Corporation, Bound Brook, N. J.

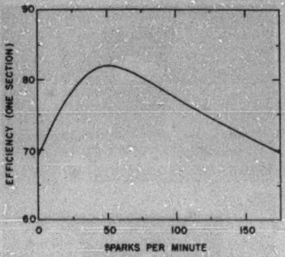


Fig. 1. Per-cent efficiency as a function of average sparking rate for a typical fly-ash precipitator

tests show that, contrary to previously accepted ideas, maximum collection efficiency usually occurs at operating voltages high enough to cause an appreciable amount of precipitator sparking. Hence, an automatic voltage regulator which is controlled by the precipitator sparking rate, rather than by one or more of the parameters mentioned in the foregoing, seems best suited to the problem.

The optimum sparking rate will vary somewhat for different types of precipitators and is best determined by actual test under the particular conditions involved. In general, however, this optimum figure typically lies between 10 and 100 sparks per minute per precipitator section. For cases involving high resistivity materials ($\rho > 10^{10}$ ohm-centimeters), the figure may be several hundred per minute. Fig. 1 shows a typical example of the effect of the sparking rate on the collection efficiency of one section of a large fly-ash precipitator.[8] The result here is a maximum collection efficiency of 82 per cent at an average of about 50 sparks per minute, dropping off to only 70 per cent efficiency both at the onset value of one spark per minute and at the excessive rate of about 175 per minute.

The sparking potential in an electrostatic precipitator is a complicated function of many factors such as voltage wave shape, electrode geometry, gas characteristics, properties of the collected material, particularly its electrical conductivity, and the amount of material which is allowed to accumulate on the electrodes. Sparking may occur at almost any value of load current over a range of voltages from about 30 to 80 kv, depending on the particular application and type of precipitator employed. It is generally random throughout the precipitator, and the occurrence of any given spark cannot be predicted.

However, at any given time, the average rate of sparking integrated over a reasonably long time, such as 1 minute or more, may be correlated with the operating peak voltage. The result for a typical fly-ash precipitator is shown in Fig. 2. The exponential increase in average sparking rate with peak voltage is an empirical relationship which has been found typical of many industrial precipitators. In any given application, the slope of the curve and the voltage intercept corresponding to an average of one spark per minute will vary somewhat with the various precipitator load and operating conditions. The resistivity of the material, the corona current density, and the voltage wave shape are among the important parameters known to affect both the slope and the voltage intercept. With collected materials in and above the critical resistivity range (order 10^{10} ohm-cm), the threshold voltage may be substantially reduced. In such cases, the slope of the curve generally is also lower than that with more conducting materials. As a result, the optimum sparking rate for best efficiency may be several times the typical figure of less than 100 sparks per minute.

Referring to Fig. 1, it is seen that the increase in efficiency from about 70 per cent at a sparking rate of one per minute to about 82 per cent at a sparking rate of about 50 per minute corresponds to increasing the precipitator voltage from 45 to about 49 kv for this particular precipitator section. In numerous cases, the performance of field precipitators has been substantially improved by adjusting the operating voltage in accordance with these principles. The importance of operating an electrostatic precipitator at the highest possible voltage compatible with the losses resulting from sparking cannot be overemphasized.

The critical dependence of collection efficiency on the operating peak voltage as illustrated in Figs. 1 and 2 is characteristic of electrical precipitators. On account of the steep voltage-current characteristic of the corona discharge, small changes in voltage can produce large changes in current and in power. This is especially significant in the region of the sparking potential where the corona power increases at a rate at least four or five times faster than the peak voltage. High voltage and high corona power promote high collection efficiency.[4,5] On the other hand, excessive voltage causes a rapidly increasing sparking rate which reduces the efficiency. Obviously, there exists a balance between these opposing factors where the gains obtained by increasing the voltage are offset by the losses resulting from sparking.

From quantitative field measurements under actual operating conditions, we conclude that the basic requirements for an automatic voltage control system suitable for a wide variety of industrial precipitator applications are as follows:

1. Operation should be controlled by the average sparking rate which should be adjustable within wide limits.
2. The sparking rate must be integrated over a reasonably long time (i.e., 1 minute or more) to obtain an average rate not greatly influenced by occasional bursts of sparking sometimes caused by bits of falling dust, momentary process fluctuations, etc.
3. The precipitator peak voltage must be smoothly adjustable to within 1 or 2 per cent of normal operating voltages. This generally means being able to set operating values within about ±0.5 kv.

Reliability, simplicity, ease of adjustment, and reasonable cost are additional requirements which are of great practical importance.

Description and Operation

GENERAL

Conventionally, an electrostatic precipitator is energized by a unidirectional voltage of negative polarity derived from the rectified output of a high-voltage alternating-current transformer. The precipitator operating voltage is adjusted by varying the transformer primary voltage by means of an appropriate regulator connected to the line source. A suitable current-limiting device, such as a series resistor, is used to attenuate primary current transients caused by precipitator sparking.

The precipitator automatic voltage control method is based on measuring the average sparking rate, comparing the measured rate with the desired optimum rate, and using the difference or error

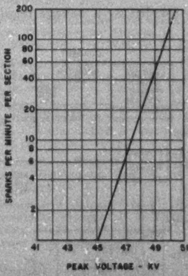


Fig. 2. Variation in sparking rate with operating peak voltage for a representative fly-ash precipitator

signal to either raise or lower the precipitator voltage to compensate for any change in the sparking rate. It may be noted that such a system inherently provides automatic adjustment for all normal operating factors, including line voltage fluctuations, which tend to alter the average sparking rate in the precipitator.

The first apparatus to regulate the voltage according to these basic principles was installed on an industrial precipitator in this country in 1949. Using a mechanical integrator with periodic measurement of average sparking rate required a minimum of about 5 minutes between precipitator voltage adjustments. In many cases this is entirely satisfactory and, indeed, several of these systems are currently in operation. The improved system described in this paper was developed: 1. to provide continuous sparking rate measurement permitting automatic voltage control (avc) operation in a greater number of applications, 2. to simplify field adjustment of control range, and 3. to eliminate manufacturing and maintenance problems associated with mechanical devices.

Fig. 3 shows a schematic block diagram of the arrangement used. In this system the precipitator average sparking rate is converted to a proportional d-c current by means of a thyratron pulser circuit and integrating network. The thyratron is triggered by transient voltage pulses caused by precipitator sparking. The indicating current meter in the integrating circuit may be calibrated directly in sparks per minute, and is of the contact-making type having two adjustable contact pointers. These pointers may be easily and quickly set to the desired optimum sparking rate range anywhere on the meter scale. Appropriate relay servo circuits responsive to the meter contact closures operate a small reversible electric motor driving a voltage regulator in the primary circuit of the precipitator rectifier set. The precipitator voltage is thus automatically controlled to maintain the predetermined optimum sparking rate required for maximum collection efficiency. The precipitator voltage adjustment may be either on a continuous or an intermittent basis, the latter regulated by a timer.

With the exception of the capacitance-type spark senser located in the precipitator high-tension circuit and the motor-driven primary voltage regulator, the automatic voltage control components comprising a single-tube amplifier stage, electronic integrator and rate meter, associated low-voltage power supplies, and motor control elements may be assembled in a compact unit of over-all dimensions 18- by 10- by 10-inches as shown in Fig. 4.

All the components are conservatively rated for maximum reliability and long life under continuous-duty conditions. The contact-making meter, recently developed by the Weston Company, is especially designed for industrial service. The thyratron is filled with Xenon gas so that its operation is unaffected by normal ambient temperatures encountered.

The important characteristics and design features of the system may be summarized as follows:

1. Continuous indication of precipitator average sparking rate on a meter scale which may be calibrated directly in sparks per minute.

2. Electronic integrator uses no moving parts and thereby avoids critical manufacturing problems and the effects of wear which are inherent in mechanical devices.

3. Sparking rate control range may be easily adjusted in the field by simply turning two pointer knobs on the front of the contact-making meter.

4. May be used with the double half-wave rectifier circuit energizing two precipitator sections from a single electrical set. In this case, two capacitance-type spark detectors are used with a selector switch which permits average sparking rate measurement in each section separately, or in both sections simultaneously.

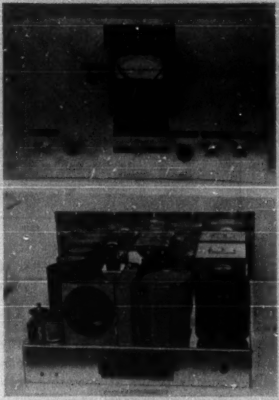


Fig. 4. Experimental model of the automatic voltage control unit

5. May be applied equally well to existing rectifier set installations and to new precipitator energization equipment, including the recently developed pulse system.[6]

CIRCUIT OPERATION, SEE FIG. 3

When an interelectrode spark occurs, the precipitator voltage collapses momentarily to almost zero. This steep wave front produces a positive voltage pulse in the capacitance divider which is composed of the pick-up element and the capacitance of the shielded input cable. Normal a-c line frequency variations in the signal are eliminated in a differentiating circuit, and the resulting voltage spike is fed through a gain control potentiometer and a double-triode amplifier stage to the grid of the thyratron. Each trigger pulse arriving at the grid causes the tube to fire and to discharge capacitor C_p through inductance L in the anode circuit. The inductance is used to limit the peak current through the thyratron and to ensure rapid deionization following the pulse. After each discharge, capacitor C_p is recharged from a d-c supply through resistor R_1 and the network composed of R_2, R, C, and the meter. The circuit equations for a pulser of this type are well known. For our purposes, it is sufficient to note that it constitutes a very simple means of producing a standard current pulse. The average value of a series of such pulses is a measure of their mean rate, and hence of the precipitator average sparking rate. The conversion of a series of events into standard pulses which are applied to an integrating circuit whose

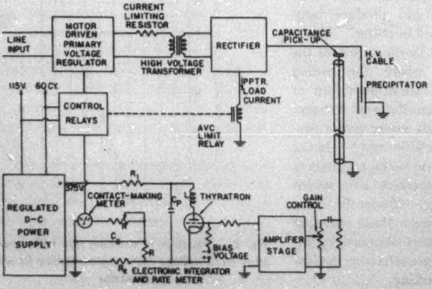


Fig. 3 (left). Schematic block diagram of automatic voltage control system for electrical precipitators

output is a measure of the average rate of occurrence of the events has found extensive application in the past 20 years. (Smith reviews basic types of pulse-shaping circuits, discusses operating theory of several integrating circuits including effects of random pulse input, and gives references to other pertinent papers.)[7]

The average current in the charging circuit is given by

$$I_c = C_p(V_c - V_0)f$$

where

f = recurrence rate, pulses per second
V_c = initial voltage across C_p
V_0 = voltage across C_p at the end of the discharge

As long as C_p and $(V_c - V_0)$ are constant, the average current will be proportional to the pulse rate. These conditions can be sufficiently satisfied by making C_p fixed, allowing adequate charging time between pulses, and using a stabilized d-c power supply of conventional type. In practice, to obtain adequate integration and smoothing, R_2 is made much smaller that R so that only a small portion of the total current passes through the indicating meter. A proper response of the control circuit to variations in the sparking rate requires at least 20 or 30 seconds for the integrating time constant. Resistor R in series with the meter is adjustable for calibrating purposes.

Multiple meter scales may be obtained by using different values of capacitance C_p with a selector switch. The unit in Fig. 4 has two scales, 0–500 sparks per minute and 0–2,000 sparks per minute.

Performance and Application

An experimental avc system of the type described was installed about a year and a half ago to control the high-voltage output of a 15-kva rectifier set energizing two sections of a large multisection fly-ash precipitator. The system has been in continuous and satisfactory service during this period without adjustment or operating difficulties. As in many fly-ash precipitators, the electrical operating conditions varied over a wide range depending on the boiler load and the type of coal being burned in the furnace. Performance was checked over a typical 14-day period by a recording milliammeter indicating the average load current of the rectifier set. While maintaining an average precipitator sparking rate between 25 and 75 sparks per minute per section, the total load current varied between extremes of 90 ma and 210 milliamperes, the latter figure corresponding to approximately full-power output from the rectifier set. The daily average current was about 180 milliamperes, varying, with periods of several hours at steady load, between 160 and 200 milliamperes. Under similar conditions normal practice without automatic voltage control would provide a maximum of about 160 milliamperes. During this particular period, the avc action resulted in an average corona power increase of at least 15 per cent above that which could otherwise be expected without frequent periodic manual adjustment by an operator.

Under favorable conditions, full rectifier power output may often be applied with little or no precipitator sparking. In such cases, the avc limit relay shown in Fig. 3 opens part of the automatic voltage control circuit so that the equipment cannot be operated above its rated capacity. As long as such favorable conditions exist, the precipitator operating voltage will be maintained at that value corresponding to full rated load current for the rectifier set. If, at some later time, the operating conditions should change so that excessive sparking occurs, the automatic system resumes control and makes the appropriate correction.

It may be well to point out that an automatic voltage control system is not a substitute for normal protective devices such as overload relays or proper current-limiting impedance in the precipitator electrical equipment. On the other hand, with avc it is important that overload relays be properly set so that full rated load may be drawn continuously. In this way, maximum utility of installed rectifier set capacity is ensured. The problem of reduced electrode clearances or a leakage path gradually developing in the high-tension circuit such as might be caused by full hoppers, for example, is of interest. With an avc system, the transformer primary input voltage will be correspondingly reduced according to the increasing degree of the fault until either a voltage can be reached just sufficient to sustain full-load current, or a short circuit develops and the overload relays take the set off the line. Advance warning of a serious fault or impending kick-out in such circumstances may be given by an alarm circuit actuated if the transformer voltage dropped below a predetermined value.

Recent developments in the system include an improved spark senser and discriminator for connection in the low-voltage primary circuit of the rectifier set. This eliminates the need for long cable runs or special protective devices associated with spark sensing in the high-tension circuit. Although the equipment described used a contact-making meter as the control element, it is apparent that the basic system permits equally well the use of other control means such as, for example, a magnetic-amplifier-type servo arrangement.

Based on the reliability and performance achieved to date, the new avc system should constitute a valuable adjunct to modern equipment for energizing electrical precipitators. It is primarily intended for applications subject to variable load conditions where maximum collection efficiency can be automatically maintained by continuously operating the precipitator at the highest possible voltage compatible with the dust losses from sparking. When little or no sparking occurs, the system will automatically maintain operation at full rated power output from the precipitator electrical equipment, thereby ensuring maximum utility of installed transformer capacity. Continuous measurement of the average sparking rate shown directly on a meter gives a convenient and effective means for quickly checking over-all performance.

The equipment may be applied both to new rectifier sets and as an adjunct to existing rectifier sets. Important applications for automatic voltage control include, for example, the collection of fly ash, cupola furnace dust, cement dust, and dust from ore-sintering operations.

References

1. Some Developments in the Automatic Control of Electrical Precipitators, H. E. Corbitt. *General Electric Review*, Schenectady, N. Y., vol. 41, no. 4, April 1938, pp. 187–90.

2. Torrlikriktare för Hoga Spänningar (Dry Rectifiers for High-Voltage Applications), Algot Arvidsson. *ASEA Journal*, Vasteras, Sweden, 1948, ASEA Reg. *4909 DK621.314.63*.

3. Verfahren zur selbsttätigen Dauerregelung der Betriebshöchstspannung von elektrischen Gasreinigungs-oder-prüfapparaten. Siemens Schukertwerke A. G. Patent No. *644,756*, Jan. 24, 1930. Also other German patents involving modifications thereof.

4. The Role of Corona Discharge in the Electrical Precipitation Process, H. J. White. *Electrical Engineering*, Jan. 1952, pp. 67–73.

5. Air Pollution Prevention in Electric Generating Stations, H. A. Bauman. *Electrical Engineering*, vol. 72, no. 3, Mar. 1953, pp. 200–04.

6. A Pulse Method for Supplying High-Voltage Power for Electrostatic Precipitation, H. J. White. *AIEE Transactions*, vol. 71, pt. I, 1952, pp. 326–30.

7. Counting Rate Meters, D. G. Smith. *Electronic Engineering*, London, England, vol. 24, no. 287, Jan. 1952, pp. 14–18.

8. Electrostatic Precipitators for Electric Generating Stations, H. J. White. *AIEE Transactions*, vol. 72, pt. III, April 1953, pp. 229–42.

◆

No Discussion

A Radio Relay Remote-Control System for Frequency-Modulation Broadcast Stations

T. R. HUMPHREY
NONMEMBER AIEE

RURAL Radio Network, Ithaca, N.Y., operates a chain of 13 frequency-modulation (FM) stations on a radio relay basis. Six of these stations are owned and operated by this network. In common with most FM broadcast stations, the transmitters are located in mountainous terrain with relatively poor access, especially in the winter season.

One site in particular, located at Bristol Center, N.Y. (about 30 miles south of Rochester) posed a serious enough access problem to warrant purchase of a special tractor-ski vehicle to provide transportation for personnel and supplies during winter months.

Although the economy which successful remote control operation would afford was obvious, it was principally because of the hazard to personnel that the first investigations of the possibilities of remote operation were made.

In November 1949, Rural Radio Network petitioned the Federal Communications Commission for special experimental authority to operate FM broadcast station WVBT, Bristol Center, N.Y., by remote control from FM broadcast station WVCN, DeRuyter, N.Y. These stations are approximately 80 miles apart.

The particular path was chosen for the initial experimentation both because of the extreme distance involved and because successful operation would minimize the personnel hazard at the remote station. Solution of the problems on the long path obviously would make extension of the system for shorter distances practical.

The Commission made a grant for such operation on an experimental basis in June 1950. In making the grant, the Commission emphasized that no reduction in equipment performance tolerances would be allowed. These requirements include:

1. Metering and maintenance of power amplifier plate voltage, power amplifier plate current, and radio frequency (RF) output voltage.
2. Observation of tower light condition.
3. Continuous monitoring of center frequency deviation, modulation percentage, and aural condition of the transmitted signal.
4. Station identification announcements. These announcements must be made only on the carrier of the station identified.

In addition to these Commission requirements, operational demands required metering and maintenance of a-c line voltages, because of exceedingly poor line voltage regulation in the rural transmitter location, and control of an auxiliary power plant. Instantaneous selection of various relay receivers was necessary. Experience with the transmitters used indicated that motor control of the power amplifier plate tuning and output link coupling adjustment was required.

Although the proposed remote control operation was the first granted by the Commission and there was no indication that permanent authority would finally be issued, it was decided to base design of the original equipment on the premise that the system might eventually be expanded to include more than one station. In practice, following adoption by the Commission of regulations authorizing remote control on a permanent basis, two remote stations are being operated from a single control point and work is in progress to incorporate two additional remote stations into the system.

The broadcasting operations of Rural Radio Network depend on the retransmission of the programs received from adjacent stations. Consequently no wire circuits exist between stations. In some areas no wire facilities of any kind are available; therefore consideration was given only to control systems which would be applicable to the radio relay system used for normal broadcasting purposes. The expense of microwave equipment precluded its use for the remote control operation. Experience with multiplex facsimile transmission on the network had proved that tones in the ultrasonic range could be applied to the FM broadcast transmitters simultaneously with normal program material without deterioration of program quality.

An analysis of the functions to be performed showed them to fall into two categories: 1. those associated with the audio input requirements for normal program service for the remote station, specifically, selection of the proper relay receiver output and station identification announcements, and 2. those associated with the metering of the remote transmitter and the maintenance of operating constants by mechanical adjustment. The audio input selection function had to be independent of and not disturbed by interrogation or adjustment operations.

A dual dial system was chosen as the method to perform the job with a minimum of human error. One dial is used to select a specific bank. Dialling within the selected bank is performed by a second dial. With this system it is possible to select the bank assigned to receiver switching, using the first bank, make the proper receiver selection on the second dial; a bank assigned to telemetering can then be selected by the first dial and the desired telemetering position selected with the second dial.

Standard 10-position stepper relays were used. Six frequencies, equally spaced from 19 through 24 kc per second, were chosen. Fig. 1 shows a block diagram of the control equipment. The source tones are provided by six continuously running oscillators, the outputs of which are isolated from the keying circuits by conventional cathode follower stages. The outputs of these stages are normally grounded through dials, push buttons or keys. When the ground connection is removed from an individual channel by operation of the appropriate dial, reset push button or control key, cathode follower output is fed to the mixer stage and amplified. The amplifier output is bridged to the audio input terminals of the control station FM broadcast transmitter. The tone level is adjusted to modulate the FM carrier about 5 per cent, which corresponds to a level approximately 26 decibels below 100-per-cent modulation. Levels for each channel are equalized by gain controls located ahead of the mixer stage. Provision is made for pretransmission observation of tone levels.

To assist the control operator, a slave

Paper 54-167, recommended by the AIEE Television and Aural Broadcasting Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing December 28, 1953.

T. R. HUMPHREY is with the Rural Radio Network, Ithaca, N.Y.

The magnetic-tape play-back machine used is a product of Mohawk Business Machines Corporation.

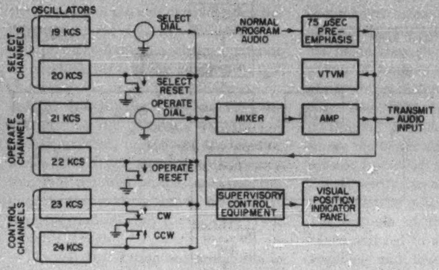

OSCILLATORS
SELECT CHANNELS
19 KCS
20 KCS
OPERATE CHANNELS
21 KCS
22 KCS
CONTROL CHANNELS
23 KCS
24 KCS
SELECT DIAL
SELECT RESET
OPERATE DIAL
OPERATE RESET
CW
CCW
NORMAL PROGRAM AUDIO
75 μSEC PRE-EMPHASIS
VTVM
MIXER
AMP
TRANSMIT AUDIO INPUT
SUPERVISORY CONTROL EQUIPMENT
VISUAL POSITION INDICATOR PANEL

Fig. 1. Block diagram of control station tone-generating equipment and supervisory control panel

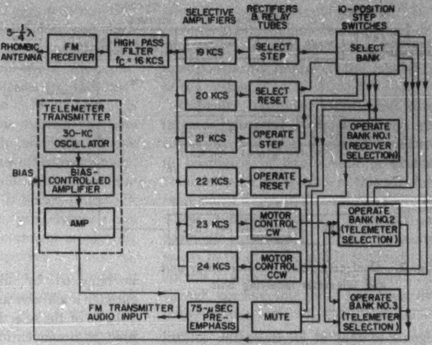

5-¼λ RHOMBIC ANTENNA
FM RECEIVER
HIGH PASS FILTER fc = 16 KCS
SELECTIVE AMPLIFIERS
19 KCS
20 KCS
21 KCS
22 KCS.
23 KCS
24 KCS
RECTIFIERS & RELAY TUBES
SELECT STEP
SELECT RESET
OPERATE STEP
OPERATE RESET
MOTOR CONTROL CW
MOTOR CONTROL CCW
10-POSITION STEP SWITCHES
SELECT BANK
OPERATE BANK NO.1 (RECEIVER SELECTION)
OPERATE BANK NO.2 (TELEMETER SELECTION)
OPERATE BANK NO.3 (TELEMETER SELECTION)
TELEMETER TRANSMITTER
30-KC OSCILLATOR
BIAS-CONTROLLED AMPLIFIER
AMP
BIAS
FM TRANSMITTER AUDIO INPUT
75-μSEC PRE-EMPHASIS
MUTE

Fig. 2. Block diagram of remote terminal equipment

unit located at the control position is driven by the foregoing equipment. The slave is similar in design to the receiving equipment at the remote station. This unit serves several purposes. Its primary function is to operate a visual position indicator panel which shows the condition of the system at any given moment. Failure of any of the control equipment is readily shown on the panel which expedites the location of trouble. Network operation requires simultaneous receiver switching at all points. The slave unit stepper performs the receiver selection for the control station.

A block diagram of the remote terminal equipment is shown in Fig. 2. A crystal-controlled FM receiver tuned to the control station frequency furnishes program material as well as the ultrasonic control tones. These tones are separated from program material at the discriminator output by a high-pass filter connected ahead of the conventional 75-microsecond de-emphasis network in the receiver. The tones are separated from one another by selective amplifier stages. The selectivity is provided by simple high-Q inductance capacitance circuits adjusted to the desired channel frequency. The output of each channel amplifier is rectified, and the derived positive direct voltage is applied as bias to a keyer tube which energizes a sensitive plate relay. The contacts of the sensitive relays are used to operate 10-position stepper relays or control the various motor functions.

One bank, or ten positions are assigned to receiver selection and station identification announcement. Two banks, or twenty positions are provided for telemetering interrogation and/or control functions. In the case of the receiver selection bank, automatic reset is incorporated. When the tenth stepper position is reached, the reset coil for the stepper is energized and the stepper returns to its off position. Simultaneously, a time delay circuit is triggered, which initiates operation of a magnetic-tape playback machine which utilizes an endless tape. Station identification announcements are recorded on the tape. Conductive paint coatings placed at 10-second intervals stop the machine automatically through wiper contacts. The announcements are spaced within the 10-second intervals. Upon completion of the interval, the machine is in a condition to repeat the cycle. Identical installations are made at both the remote and control points. To make a station identification announcement on both stations, it is only necessary to dial a zero in the receiver selection bank. The announcement is made for each station and the stepper is returned to off in preparation for the next receiver selection.

Since a number of receivers are fed to the receiver stepper switch, stepping or resetting the bank causes the stepper rotor contacts to pass over switch positions to which these "live" receiver outputs are connected. To avoid bursts of program material from undesired sources or the completion of feedback loops within the network system, electronic muting of the transmitter input is performed whenever a step or reset coil is energized in the receiver selection bank.

The telemetering steppers allow selection of sampling voltages at the remote point. The stepper contacts are interwired to allow meter observation during mechanical adjustment of pertinent controls. For example, power amplifier plate current can be observed in two positions of the stepper. In one case, this reading may be observed while the power amplifier plate tank is adjusted. The same reading is available while the output coupling link is mechanically adjusted. Only the mechanical adjustment directly affecting the meter position dialled can be performed.

The 19- and 20-kc channels respectively are used to step and reset the select stepper. This stepper, as the name implies, is used to select the desired bank in which an operation is to be performed. The 21 and 22 kc channels are used to step and reset the operate stepper. This stepper performs the desired operation within a selected bank and is controlled by a separate dial and reset push button. The 23- and 24-kc channels are used for motor control and on/off purposes and are transmitted by operation of a lever type key.

As an example, assume we wish to measure power amplifier plate current and retune the plate tank of the power amplifier stage. Assume the power amplifier plate current-plate tank adjust position is position 4 in the third bank. We would first dial a 3 on the select dial. This would transmit three pulses of 19-kc tone from the control station superimposed on the FM broadcast carrier. These pulses would be removed from the FM receiver at the remote point and separated as previously explained. The three pulses of 19-kc energy would be rectified, advancing through the keyer tube the select stepper to its number 3 position. This presets the third bank for dialling within it. On the second dial at the control point, a 4 is dialled which transmits four pulses of 21-kc energy and in a manner similar to that just described, advances the stepper assigned to the third bank at the remote point into position 4. It should be remembered that, coincident with these stepper operations at the remote point, the visual position indicator panel at the control point is being keyed by the same

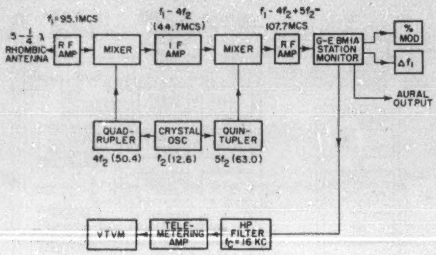


Fig. 3. Dual conversion monitor driver unit and telemeter receiving unit

signals so that at all times the control operator is aware of the operations he has performed.

When the stepper has reached the fourth position of the third bank as outlined, two circumstances exist as a result of interwiring. First, the telemeter transmitter, details of which follow, is turned on and the appropriate sampling voltage is applied to it for transmission back to the control point and, secondly, circuits are completed through interwiring to allow motor control of the power amplifier plate-tuning control. At the control point, the motor control key is thrown either to the left or right and the correct motor at the remote point will be energized as long as the key is depressed. Power amplifier plate current is observed during this tuning period and the desired effect can be achieved.

Telemetering information is returned to the control point by superimposition of a 30-kc tone on the carrier of the remote FM broadcast transmitter in a manner identical to that used to transmit the tones from the control transmitter. The telemeter transmitter consists of an oscillator driving a variable gain amplifier. The output of the amplifier stage is proportional to the value of sampling voltage applied as grid bias to the stage. The standard reference voltage used for circuit calibration is furnished by a 12-volt "hot-shot" battery. Voltages for normal operating values are derived from the various sources which are telemetered and are compared to the reference voltage. Telemeter output level is adjusted to modulate the remote FM broadcast transmitter 5 per cent when standard reference voltage or normal sampling voltage is applied as bias.

At the control point the 30-kc voltage is separated from normal program material at the discriminator output from the General Electric (G-E) *BM-1-A* station monitor. The telemetering voltage is removed by filtering and feeding through a selective amplifier. The output voltage of the selective amplifier is measured on a conventional broadcast type-*VU* meter, for which direct-reading scales have been prepared to indicate radio-frequency output voltage, power amplifier plate voltage, power amplifier plate current, and a-c line voltage. In the region of normal operating values these scales are greatly expanded. This effect provides closer tolerances at the control point than are possible at the remote point on standard meter scales.

It is of course essential that the control station maintain full control of the remote station under all conditions. To assure maintenance of this condition the filament circuits of the remote transmitter are interlocked with a carrier-operated relay circuit operated by second limiter grid voltage in the FM relay control receiver. Failure of the control station, the control receiver at the remote point, or of any of the tubes in the filament control circuit automatically removes the remote transmitter from the air. In addition, a limit control was incorporated which operates on variations in field intensity of the control station signal at the remote point. This control is adjusted to remove the remote carrier from the air if, as a result of propagation effects, the control carrier should drop below a prescribed level. This avoids the possibility of capture of the remote transmitter by another station operating on the control station frequency and whose signal might be receivable during unusual propagation periods. Discrimination against undesired signals is further abetted by the use of a multiwave-length terminated rhombic receiving antenna at the remote point. This antenna was designed for maximum front-to-back ratio rather than maximum forward gain. Over a period of 18 months, during periods of extreme propagation variations, no difficulties have been encountered on this score in spite of the relatively long path involved.

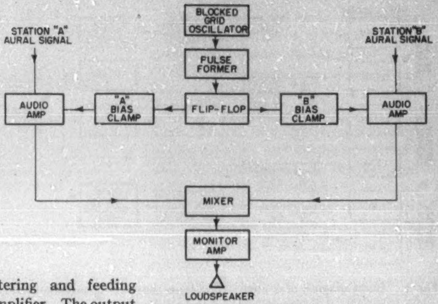


Fig. 4. Electronic switching arrangement for aural monitoring of control and remote stations

In fulfilling Commission requirements relative to center frequency deviation and modulation percentage, it was deemed highly preferable to utilize equipment which carried Federal Communications Commission approval rather than to become involved in any system which would require Commission type approval before being acceptable. As a result, the dual conversion system shown in Fig. 3 was used.

The remote station operates on 95.1 megacycles (mc). The G-E *BM-1-A* station monitor used was calibrated for a frequency of 107.7 mc, a difference of 12.6 mc. A temperature-controlled crystal oscillator operating on this difference frequency is the source for oscillator injection voltages for the dual conversion. Adjacent harmonics, the fourth and fifth, are used for injection frequencies. The fourth harmonic, 50.4 mc, beat against the 95.1-mc signal frequency in the first mixer produces a difference intermediate frequency of 44.7 mc. This frequency, when heterodyned with the fifth harmonic of the difference oscillator on 63.0 mc, produces a sum beat of 107.7 mc, the frequency for which the G-E approved monitor was calibrated. With this system, three frequencies are available at which considerable amplification can be had without interaction; the signal frequency of 95.1 mc, the intermediate frequency of 44.7 mc and the output frequency of 107.7 mc. Adequate gain is provided to provide the watt or so of radio-frequency power required to drive the G-E monitor. The monitor provides direct-reading values of center frequency deviation and modulation percentage. In addition, the monitor is used as the source at the control from which the telemetering voltage is derived and from which aural monitoring is fed.

In the conversion system just described, the accuracy of the unit depends upon the tolerance of the difference-frequency oscillator. The frequency of this oscillator is checked periodically against WWV transmissions, using a secondary standard for interpolation. In practice, the stability of this oscillator is excellent enough so that, in view of the plus or minus 2,000-cycle-per-second frequency tolerance required at the operating frequency, the error in actual transmitter frequency indicated by the approved G-E monitor is slight.

The various stations of Rural Radio Network carry identical program material with the exception of station identification announcements. Although it would be possible to furnish separate monitoring facilities for the control and remote station, loss of one station might not be immediately apparent to the operator on duty. Deterioration of program quality would be masked by the loudspeaker signal of the other signal. To make readily apparent the loss of or deterioration of either signal, the automatic monitor-switching method diagrammed in Fig. 4 was devised.

A low-frequency blocked grid oscillator furnished the switching rate time base. The saw-tooth output is shaped and fed to a conventional flip-flop circuit. To each half of the flip-flop circuit are connected clamp keyer tubes. These are sharp-cutoff pentodes, the cathode returns of which are made through the cathode bias resistors of medium-mu triode audio amplifier stages. Constants are such that when the clamp tube cathode current is maximum, adequate voltage appears across the triode audio amplifier cathode resistor to cut off this stage completely. Consequently, one audio channel is operative while the other is biased to cut off. The outputs of the two channels are mixed and fed to a common monitoring amplifier and loudspeaker. Thus, although the output signal is continuous, it is composed of equal segments from each station's program. Loss or deterioration of either signal is immediately apparent to the control operator. A sampling rate of about one sample each 6 seconds is found to be satisfactory.

Although a change in transmitter power with a resultant shakedown period interferes with precise analysis of the reliability of the remote control system, such an analysis was made. This was based on identical 4-month periods, before and after remote operation was installed. A percentage of performance figure was arrived at and it showed a reduction of 0.35 per cent. This period also involves the earliest period during which unattended operation was begun and represents the poorest records on hand. Improvements in equipment design and development of maintenance techniques have constantly improved the reliability of the system.

Difficulty has been experienced with intermittent false stepping of the receiver selection bank, apparently because of transients in live speech programs, Oddly enough, no trouble of this sort has been encountered with musical programs where difficulty was anticipated from harmonic generation. The difficulty seems to be sporadic and is confined to a specific program and one individual. Although considerable time has been devoted to observation of selective amplifier outputs during these programs, to date nothing concrete has been isolated. New gating circuits are being incorporated to eliminate this trouble. A 29-kc tone, amplitude-modulated by a 500-cycle signal will be transmitted prior to any of the operational tones. The gate envelope is demodulated and the derived 500-cycle signal is used to control the main relay supply source. Bench tests indicate that this refinement will completely cure the trouble.

◆

No Discussion

A Transducer Using a Short-Circuit Rotor

V. A. ORLANDO
NONMEMBER AIEE

Synopsis: An a-c phase-sensitive signal generator for use in servo and instrumentation work is described. The voltage gradient and zero position are independent of linear rotor motions, and the torque and magnetic side pull on the rotor are extremely small.

AS CONTROL equipment becomes more automatic and accuracy requirements increase, more highly specialized position-sensing equipment becomes necessary. Electrical position-sensing elements can be broadly separated into two classifications: position-repeating systems, and signal generators. The position-repeating systems are comprised of a transmitting unit and a remotely located receiving unit which are electrically related so that a rotation of the transmitter shaft causes an equal or proportional rotation of the receiver shaft. A large number of devices for performing this function are described in the literature.[1]

In the signal generator type of position-sensing element, sometimes referred to as a pickoff or pickup, angular rotation of the input shaft causes a proportionate voltage to appear as the output of the device. In many applications, the output is used in conjunction with electronic equipment, in which case it is usually desirable that the output be of line frequency and phase-sensitive, that is, that the voltage increase on either side of a minimum or null voltage position and the phase of the output shift 180 degrees as the rotor passes through this null position. A typical application for such a signal generator is in a servo system in which the output element is a servomotor. The torque produced by these motors depends on the amplitude of the signal generator voltage and the direction of rotation is determined by the phase of the signal voltage. This paper describes a signal generator which is particularly adapted to instrumentation and laboratory application although it is not limited to these fields.

A number of different types of signal generators are available which operate on a variety of principles.[2] Those most widely used are variable reluctance-induction devices characterized by a stator which carry the windings and a soft iron rotor or armature which completes the magnetic circuit of the stator. The a-c excitation flux is carried through two separate stator magnetic circuits, each circuit having an output winding. These output windings are connected in series opposition so that when the rotor is at the null position, equal magnitudes of flux pass through each of the two output windings and the resulting induced voltages add to zero.

Paper 54-172, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted October 14, 1953; made available for printing November 27, 1953.

V. A. Orlando is with the General Electric Company, West Lynn, Mass.

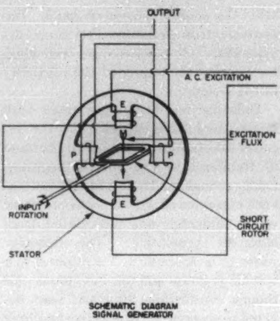


Fig. 1. Diagram showing principle of signal generator

As the rotor is rotated, increasing the reluctance of one magnetic circuit and decreasing the other, voltages in the two output coils become unequal and a net output from the signal generator results.[3]

The signal generator described in this paper has a similar stator construction to that just described except that the output windings are connected series-aiding, and no flux passes through the output windings at null. The rotor, instead of being soft iron, is a closed loop of nonmagnetic conducting material. This construction allows the signal generator to have many of the operating characteristics of the variable reluctance types mentioned and two important additional features: (1). the output is practically unaffected by linear motions of the rotor, and (2). magnetic side pull on the rotor is negligible even when it is not centered.

Principle of Operation

As shown in Fig. 1, the pickoff is comprised of a soft-iron laminated stator having four salient poles. On each pole is a coil having a simple winding. The two sets of diametrically opposite coils are wound series-aiding. One set, marked *E* acts as the excitation windings; the other, marked *P*, as output windings. A short-circuited turn of nonmagnetic conducting material, such as copper, is placed so that it links the excitation flux. With the rotor in this position the a-c excitation flux generates a large circulating current in the short-circuited turn. If the plane of the short-circuited turn is inclined slightly with respect to the axes of the output windings, a component of the flux generated by the short-circuit currents will link the output coils and an output voltage will result. Since, for small angles, the component of flux linking the pickoff coils is proportional to the angle through which the short-circuited turn rotates, the output is linear with angle. The output is phase-sensitive since the direction in which the flux passes through the output windings depends on the side of null from which the short-circuited turn is rotated.

If the short-circuited turn were rotated 90 degrees so that it linked no excitation flux, no circulating current would be generated in it. This is a second null position of the signal generator. If the short-circuited turn is rotated slightly from this position, a component of excitation flux will pass through it and a small circulating current will flow. This short-circuit current constitutes ampere-turns about the output coils. The result is that similar gradients, i.e., the derivative of output voltage with respect to angular position, are obtained about any of the four 90-degree null positions.

In instrumentation and other precise work it is often necessary that the torque exerted on the rotor of the pickoff be very small. In this signal generator the torque is zero at the null position but increases slightly with rotation. From fundamental considerations it can be shown that the rotor tends to turn so that no flux links it. In the null position where the short-circuited turn rotor links all the excitation flux, the rotor is in a position of unstable equilibrium since it tends to rotate to one of the null positions 90 degrees away. This results in what may be termed a negative spring torque, for, as the rotor angle increases from null, the torque on it increases tending to move it away from this null position. However, in the second null position where the excitation flux does not link the short-circuited turn, the torque can be termed a positive spring torque, since moving it from the null position increases the torque, tending to return it to this null position. In practice it is found that the torque gradient through the unstable null position is very small compared to the torque gradient through the stable null position. This is the reason for usually choosing to operate about the unstable null position.

From the principle of operation and from geometric considerations, it can be seen that moving the short-circuited turn rotor axially or in any radial direction will induce negligible voltages in the output winding when the rotor is at the null angular position. Similarly, geometric relationships are such that the voltage and voltage gradient are relatively independent of small linear motions of the rotor.

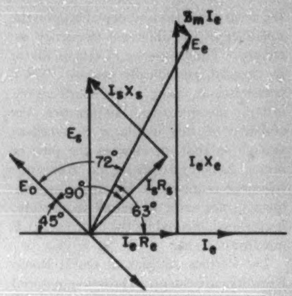


Fig. 2. Phase-angle diagram of signal generator

E_0 = output voltage
E_e, I_e = excitation voltage and current
E_s, I_s = voltage and current induced in rotor
R_e, X_e = excitation winding resistance and reactance
R_s, X_s = rotor resistance and reactance
Z_m = mutual impedance, rotor and stator

Mathematical analysis is not conveniently used to calculate the output of this signal generator because the air gaps are large and it is difficult to calculate the impedance of the short-circuit rotor. However, it has been found of great value to examine mathematical relationships because they give an accurate picture of the interactions involved between parameters and they relate and extend the usefulness of experimental data. Based on simple coupled circuit theory, the output of the signal generator may be written

$$e_0 = \frac{i_e \omega^2 M_{es} M_{so}}{\sqrt{R_s^2 + \omega^2 L_s^2}} \text{Arctan} - \frac{\omega L_s}{R_s}$$

where

e_0 = signal generator output
i_e = excitation current
ω = excitation angular frequency
M_{es} = mutual coupling, excitation windings to rotor
M_{so} = mutual coupling, rotor to output windings
R_s, L_s = rotor resistance and inductance

From this relationship it is apparent that the output varies with frequency to a power between the first and second. It is also evident that the output phase angle is equal to the negative phase angle of the rotor impedance.

The mutual coupling terms are highly significant. One term, depending upon which null position is used, is very nearly constant with output, but the variation of the other with rotor angle determines the output of the signal generator. It can be shown that this mutual term changes sign as the rotor passes through the null position, and increases until the rotor angle exceeds half the angle whose cord is equal to a stator pole width. Beyond this angle the magnitude of the mutual term tends to remain constant except for the rather large effect of the flux leakages involved. The phase-angle diagram shown in Fig. 2 relates the significant voltages and currents.

Description of Components

A practical embodiment of the principle described here is shown in Fig. 3. This form of the signal generator is comprised of stator, rotor, and null adjuster, all of which are important in increasing the ratio of the output voltage gradient to null voltage. Since this ratio is a basic measure of merit for a signal generator, the means of increasing it are of importance.

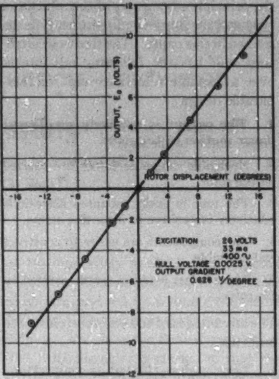


Fig. 4. Performance curve, signal generator

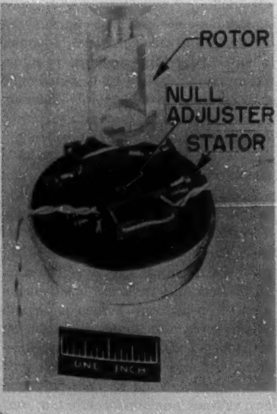


Fig. 3. Signal generator with short-circuited turn rotor. Rotor removed from stator

When optimum performance is required, a null adjuster is used. In Fig. 3 it is shown as a nearly circular laminated-iron assembly centrally located in the gap of the 4-pole stator. It can be rotated with respect to the poles for adjustment purposes, and thereafter is mechanically fixed to the stator. It has two important functions; the first is to increase the output gradient by reducing the reluctance of the flux paths, the second is to reduce the null voltage. Without a null adjuster, the null voltage of the pickoff is not a minimum because of nonhomogeneity in the material and nonsymmetry in the construction of a stator. By making the null adjuster slightly asymmetric, rotating it will generate a component of voltage in the output winding. This component of voltage can be adjusted to reduce the voltage at the null position. The voltage induced in the output windings by the rotor circulating current is out of phase with the component of voltage controlled by the null adjuster and also probably out of phase with the original null voltage. By alternately adjusting the null adjuster and rotor, the resulting null voltage can be made very small and the fundamental excitation frequency can be eliminated from it. The remaining null voltage is comprised of third and higher odd harmonics of the excitation frequency and is primarily the result of nonlinearity of the magnetization curve and nonuniformities of the iron in the circuit.

Rotors of various shapes and materials have been used satisfactorily. A simple shape is merely a closed loop of material as shown in Fig. 1, or even simpler, a flat strip or slab. The rotor construction shown in Fig. 3 allows the use of a simple but effectively shaped null adjuster. Any nonmagnetic, conducting material can be used for the rotor. However, the maximum ratio of signal gradient to null voltage is obtained when a high-conductivity material is used. When extremely low torques are required a low-conductivity material should be used. A material having a low thermal coefficient of resistance may be used to minimize the variations caused by changes in temperature.

Performance Characteristics

The characteristics reported here are for typical signal generators that have rotors of high conductivity aluminum. Aluminum was used because it was desired that the rotors be of light weight. One side of both the excitation and output windings was grounded to the stator to reduce stray pickup. The output was measured on a 0.5-megohm vacuum tube voltmeter. The excitation power was supplied by an audio oscillator at 26 volts, 400 cycles, unless otherwise stated. Each unit consumed about 0.40 watt of power.

The calibration curve shown in Fig. 4 has a gradient of 0.628 volt per degree and a null voltage of 0.0025 volt, and is linear for ±10 degrees from null. By increasing the excitation current to produce a gradient of 1.0 volt per degree, some units maintained a null voltage of less than 0.003 volt.

Varying the excitation current ±10 per cent varied the gradient by ±10 per cent. A ±10-per-cent variation of frequency with constant excitation current varied the gradient ±17.0 per cent. A ±10-per-cent variation of frequency with constant excitation voltage varied the gradient by ±10 per cent.

The variations in gradient with temperature using the gradient at 25 degrees centigrade (C) as a reference, keeping the frequency and current constant were

+19.6 per cent at −70 C
+10.4 per cent at −30 C
−10.7 per cent at 98 C

During all these tests the angular position at which the null voltage occurred did not vary more than the test equipment sensitivity estimated to be 0.003 degree.

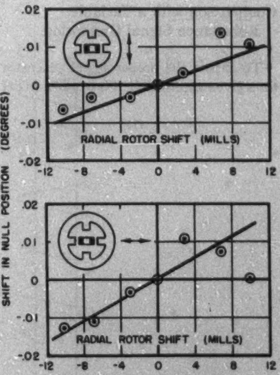


Fig. 5. Angular shift in null position with radial rotor motion

Arrows indicate directions of rotor displacement

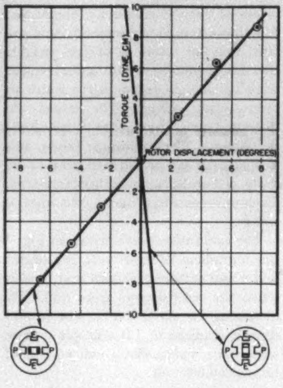


Fig. 6. Rotor reaction torque

Shifts of the null position with linear motion of the rotor are shown in Fig. 5. It is very likely that the variations could be decreased if unusual care were used in machining the rotor for flatness and symmetry. The change in gradient with linear shifts of the rotor over a ±0.015-inch distance was negligible.

The rotor reaction torque is shown in Fig. 6. The sketches below the curve indicate the rotor null positions about which the measurements were made.

The phase angle between the excitation voltage, the excitation current, and the output voltage were very nearly constant with output. The values of these angles for the units tested are shown in Fig. 2.

Comparison with a Variable Reluctance Signal Generator

To choose the proper signal generator for a given application, it is of value to list the principle advantages of the reluctance type and the short-circuited-turn type described in this paper. The short-circuited-turn unit has the following advantages over a variable reluctance unit of comparable design:

1. The output is relatively unaffected by linear motion of the rotor.
2. Negligible magnetic forces are exerted on the rotor.
3. The unit is easily mounted since it is relatively insensitive to mounting stress.
4. The components are easier to construct and assemble.
5. There is less reaction torque exerted on the rotor.

The advantages of the variable reluctance units are:

1. They are less affected by temperature.
2. They are less affected by frequency variation of excitation source.
3. Linear output is over a wider range.
4. A null adjuster is not required.
5. There is inherently higher signal gradient.

Special Applications

The stator and coil assembly of the signal generator described has many applications in a laboratory other than in the particular embodiment described here. For merely measuring the angular rotation of a shaft, it is not necessary to use a null adjuster with the stator assembly unless high gradient and a low null are required. The fact that the rotor can be made of any nonmagnetic conducting material in almost any shape allows wide latitude in improvising a rotor for special applications.

If a signal proportional to linear motions is required, a flat strip of nonmagnetic conducting material twisted in the form of a helix can be used as the rotor. This arrangement is of course sensitive to both rotory and linear axial motions of the rotor.

Occasionally, it is desirable to produce controllable torques of small magnitudes. By exciting both the excitation and output winding, a small torque can be produced on the short-circuited rotor. The torque is generated in a manner similar to that in a 2-phase servomotor, and similar controlling arrangements can be used.

Conclusion

The signal generator described in this paper has features which allow it to be incorporated in highly sensitive, accurate devices without requiring the components to be unusually accurately machined or aligned. It is very versatile in application and fulfills most of the requirements of more familiar signal generators. Although it is expected that it will find widest application in the field of instrumentation, it can be used in most electrical control systems which require a null type of sensing device. To obtain full advantages of the signal generator it is desirable to make it an integral part of the associated equipment.

References

1. ELECTRICAL REMOTE POSITION-INDICATING SYSTEMS AS APPLIED TO AIRCRAFT, **R. J. Garvey.** *Journal*, Institution of Electrical Engineers, London, England, vol. 94, pt. IIa, 1947, pp. 283–91.

2. MECHANICAL MEASUREMENTS BY ELECTRICAL METHODS (book), **H. C. Roberts.** Instruments Publishing Company, Inc., Pittsburgh, Pa., second edition, 1951, chap. IV.

3. MICROSYN ELECTROMAGNETIC COMPONENTS, **R. K. Mueller.** Instrumentation Laboratory, Massachusetts Institute of Technology, Cambridge, Mass., 1952.

◆

No Discussion

High-Gain Side-Firing Helical Antennas For Ultrahigh-Frequency Television Broadcasting

HOWARD G. SMITH
ASSOCIATE MEMBER AIEE

IN designing any high-frequency device which is to have relatively uniform characteristics over a band of frequencies, one of the prime considerations for easing the severity of design compromises is to choose a basic configuration inherently suited to maintaining the reactive energies of the system low compared with the energy flow per cycle through the device. Such a smooth flow through the system can be approached by maintaining the impedances throughout the system resistive over the desired performance frequency band. This in turn dictates the avoidance of resonant components since such resonant elements tend to introduce high reactance slopes into the system impedances requiring careful compensation in the design. The case of the uniform transmission line with distributed constants properly arranged so as to minimize reflections is perhaps the ideal example of such a system.

In the design of a radiating system which will have uniform gain at all azimuths and achieve this gain by concentration of the radiated energy within a relatively small angle in the vertical plane, it becomes necessary to illuminate a cylindrical vertical aperture by multiple sources of coherent phase at a given azimuth, while maintaining the variation of illumination of the radiating aperture at a minimum with respect to azimuth. It is desirable, moreover, to achieve this illumination of the vertical aperture in such a way as to require, if possible, a minimum number of feed points with the necessary matching and power dividing networks with attendant complications of circuitry.

The design of television broadcasting antennas for service in the ultrahigh-frequency band has as its basis both of the design objectives mentioned. Consequently, it is natural to explore possible configurations which might be contrived to maintain mechanical simplicity while achieving such an illumination of aperture, and while maintaining freedom from reactive storages. Examination of the configuration indicated in Fig. 1, consisting of two oppositely wound helices disposed concentrically about a conducting cylinder as shown, and fed by a source connected between their common junction and the cylinder, will show the desired properties. The helix diameter is chosen so that the currents in all conductors are in phase at a given azimuth, and the cylinder diameter is chosen so that the radiation of energy per turn is sufficiently large that the reflected waves from the ends of the helices are minimized, but is not chosen so large as to ineffectively illuminate the aperture required for the desired concentration in the vertical plane.

In such a configuration each helix and its associated conducting cylinder may be considered much the same in nature as a single wire transmission line using the cylindrical surface as the return path. As is well known, such a line tends to radiate energy in increasing amounts as the spacing between conductor and return surface is increased. Hence the relative diameter of the helix and the cylinder constitute an effective control on the radiation of energy per turn. This spacing can be used to accomplish a proper balance between degree of aperture illumination and standing wave ratio present on the helix. By choice of the helix diameter so that one turn of the helix constitutes a length equal to an integral number of wave lengths, the currents in all conductors of the antenna possess a coherent phase relationship at a given azimuth, and consequently radiate as a broadside array to concentrate the radiation in a vertical plane.

In the choice of methods of illumination of an aperture of coherent phase if a uniform intensity of illumination over the aperture is chosen, a highly concentrated main beam of radiation is secured but with relatively high side-lobe levels. On the other hand, a tapered illumination of the aperture gives a somewhat wider main beam width for a given physical aperture and achieves a reduction of side-lobe number and level compared with the uniformly illuminated case. For the antenna bay configuration shown, the illumination along the vertical direction will be tapered due to a decrease in current along the helices, giving a good suppression of side lobes. The stacking factor of several such bays placed in a vertical tier enables the designer to achieve further narrowing and shaping of the main beam to give the required gain for a multiple-bay array. An adjustment of radiation per turn which results in a 20-decibel attenuation due to radiation between the feed point and the helix ends, for example, even with complete reflections, yields a reflection coefficient of 1/100 back at the drive point, and hence gives a resistive driving point impedance characteristic which is essentially

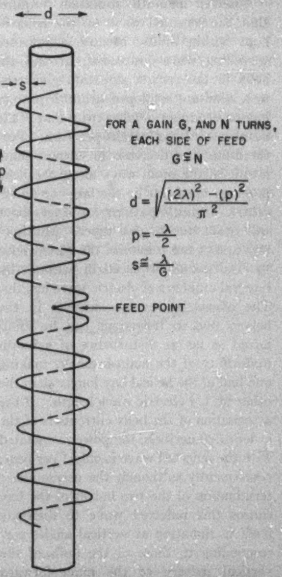

d
s
p
FOR A GAIN G, AND N TURNS, EACH SIDE OF FEED
$G \cong N$
$d = \sqrt{\frac{(2\lambda)^2 - (p)^2}{\pi^2}}$
$p = \frac{\lambda}{2}$
$s \cong \frac{\lambda}{G}$
FEED POINT

Fig. 1. Winding arrangement of single bay side firing antenna composed of concentrically disposed helical conductor and conducting cylinder

---

Paper 54-165, recommended by the AIEE Television and Aural Broadcasting Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing December 15, 1953.

H. G. Smith is with Cornell University, Ithaca, N. Y.

The author wishes to acknowledge the assistance of L. O. Krause and R. E. Fisk of the General Electric Company (with whom the original investigation of the helical antenna was undertaken in 1950) in making available for inclusion in this paper the material of Figs. 2, 3, and 4 on later-developed commercial models.

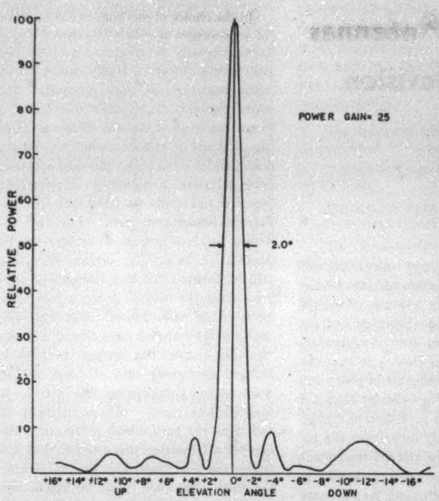


Fig. 2 (above). Measured vertical pattern, on a power basis, of 5-bay antenna without special phasing for null fill-in or beam tilt

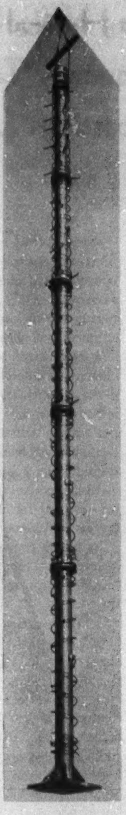


Fig. 3 (right). Assembled 5-bay helical antenna

that of a travelling wave device having uniform characteristics over a relatively broad frequency band.

Uniform azimuth radiation requires that the reflected wave on the helix be kept within limits, because a reflected travelling wave proceeding around the helix in the reverse direction gives rise to a resultant scalloped azimuth pattern having a number of lobes twice that of the number of wave lengths per turn. Furthermore, the decrease in current magnitude of the main wave along the helix per turn must not be too large or an eccentric azimuth pattern for single bays will result due to this taper. Multiple-bay arrays can minimize the eccentricity by feed points displaced in azimuth by integral numbers of electric wave lengths. The effects of reverse currents in the helices due to reflections can be minimized as far as disturbance of azimuth uniformity of the main beam by making one-half of the helical bay longer than the other by 1/4 electric wave length. If the attenuation of the helix currents is of the order of 20 decibels, the power associated with the reflected wave is only 1 per cent; consequently, although the assymmetric termination of the two halves of the bay causes this reflected wave to dissipate itself in radiation at vertical angles corresponding to those of the nulls of the vertical pattern of the main forward going waves on the helix, the amount of power so lost is negligible. This achieves the elimination of any undesired assymmetry in the main beam azimuth pattern and modifies slightly the nulls in the vertical pattern by radiating the reflected energy in these directions. Supporting insulators for the spacing of the helix from the center cylinder or supporting mast introduce certain discontinuities which, though small, would tend to give rise to slight standing waves on the helices and result in a nonuniform azimuth pattern. These discontinuities are readily compensated by the spacing of consecutive supports at odd multiples of 1/4 wave length or by relative skewing of the support points on the upper and lower halves of the bay, just as in the case of residual end reflection elimination, by 1/4 wave length.

The accurate control of the elevation of maximum radiation of the main beam within a few degrees can be readily achieved by control of the phase of the drive points of several bays stacked in a vertical tier. Uniform phase advance of the drive points of each succeeding bay causes a downward tip of the beam. With the feed of each bay being made coaxially through the mast this can be readily achieved by rotation of the drive points in azimuth without disturbing the impedance relationships between the feeds. When as many as five feed points are so arranged, several degrees of main beam tilt can be achieved with inappreciable beam breakup. In antenna designs with only two such phase control points, a beam tilt of a single degree begins to result in noticeable deterioration of the main beam. Although it is not necessary for the beam tilts usually desired, it is possible to secure even larger beam tilts without beam breakup by the use of multiple-bay stacking factors as described, and by adjusting of the turn lengths of the individual bay helices so as to cause the beam center of the individual bays to have a corresponding depression angle. This can be readily accomplished by decreasing the helix diameter in the upper half of each bay and increasing its diameter on the lower half of each bay to set up the desired phase advance progressively from the top to the bottom of each bay.

In order to achieve sufficient mechanical stiffness to warrant a relatively high concentration in the vertical plane with its attendant high gain (a beam width of 2 degrees at half power points being encountered for power gains of 25), it is usually necessary to have a cylindrical supporting mast sufficiently large in diameter so that it is impossible to use a helix with a single wave length per turn. For this reason multiple wave-length turns are normally employed. On the other hand, use of a large number of wave lengths per turn results in unnecessarily large diameters with their attendant mechanical disadvantages. Since the turns of each half of one bay are fed in series with one another, increasing the number of wave lengths per turn also causes progressive deterioration of the phase coherence of the turns with incremental frequency departure from the center-design frequency, and hence narrows the frequency range over which the antenna will retain its beaming gain characteristics. For these reasons a choice is normally made in favor of helices having a turn length of 2 wave lengths. The choice of pitch of the helix is then dictated by a compromise between a desire for a small pitch to reduce the crosspolarization components of radiation from the helix, and a desire for a large pitch to reduce the number of turns utilized in illuminating a given aperture. The latter is desirable since a large number of turns fed in series gives rise to narrowing of the beaming bandwidth due to deterioration of phase coherence at other than the design frequency. If one chooses a pitch of 1/2 wave length for a turn length of 2 wave lengths, the pitch angle is of the order of 15 degrees. This tends to reduce the cross-polarized components of radiation from elemental lengths of the helix in directions normal to its axis to small values. The winding of the two halves of each bay in the opposite sense, moreover, causes the cross-polarized fields at the elevation of the desired beam maximum of each bay to be essentially cancelled. The cross-polarized fields tend to be a maximum at the elevation of the minimums of the individual bay radiation patterns, but due to the small pitch angle and the loss of phase coherence between radiation from the individual turns at these angles, the energy so radiated has been shown by measurement to be of the order of 5 per cent of the total energy radiated.

A choice of a pitch of 1/2 wave length is such, moreover, that if the helix to mast spacing is of the order of 1/10 wave length (an appropriate spacing for single bays having a gain of 10 per bay), the mutual impedance between turns is negligible with respect to the self-impedances of the line. Under these conditions, velocity corresponding to that in the appropriate dielectric media is obtained. Choice of smaller power gain per bay requires wider mast to helix spacing for proper attenuation, and the effects of the mutual impedances between turns may modify the required helix turn length by slight modification of the phase velocity along the helix.

A further control of the vertical radiation pattern of an array is sometimes desired in order not only to depress the main beam by a specified angle below the horizon but also to insure that coverage holes do not occur close in to the transmitting tower due to the presence of severe nulls. Reduction of the severity of the nulls of a pattern can be achieved in an array by disturbing the uniformity of the phasing of the coherent sources or by alteration of the systematic illumination of the vertical aperture. In the case of the helical multibay system described, after the depression of the main bay is achieved by progressive azimuth rotation of the feeds to secure phase advance of the upper bays, a single bay can be advanced or retarded with respect to the phase system of the group by, for example, 30 degrees. At an elevation then which would have been an absolute null in a 5-bay stacked antenna, the five components of the stacking factor phasors do not add to zero, but approximately to that of the contribution one bay multiplied by the sine of 30 degrees. This gives a field at the nulls which is approximately one-tenth that on the main beam, or 20 decibels down, as a minimum. This phasing is readily achieved by bay rotation without disturbance of the impedance conditions of the array, which would result if an attempt were made to alter the power division between the vertical elements of the array in order to modify the illumination distribution. The phasing adjustment is susceptible to modification to meet conditions encountered in the field.

The actual design compromise between a large number of bays and feed points on the one hand, and a single bay capable of the desired over-all gain with a single feed, depends largely on the required bandwidth to be achieved in a given design. Also to be considered is the required uniformity of gain to be achieved over a channel group using the discreet mast sizes available to the manufacturer. If the limiting case of attempting to achieve all the required gain in a single bay is examined, it is found that for a gain of perhaps 25 using a helix having a turn length of 2 wave lengths and a pitch of 1/2 wave length, there will be a length of helix between feed and termination of 50 wave lengths. It does not take a large increment of frequency relative to the design center on such a device to cause rapid deterioration of phase coherence along the helix resulting in beam breakup. It would not be possible to secure the required 6 megacycles bandwidth even in the ultrahigh-frequency band. Furthermore, in such an extreme case thermal elongation of the helices over the normal operating temperature range might change the design center by as much as one full channel. In order to achieve a practical bandwidth, then, with enough spare to absorb thermal elongations without adverse effects, a compromise is made to increase the number of feeds from one to five. This not only solves the beaming bandwidth problem but also makes available a stacking factor due to multiple bays for beam tilt and shaping.

A plot of experimental data on a 5-bay helical design having a power gain of 25 relative to a dipole is shown in Fig. 2. The data shown are for a stacking arrangement involving no beam tilt or null fill-in phasing.

A choice of the gain per bay then which in effect divides the feeding system into five feeds in parallel, each of which in turn consists of a pair of five vertically stacked turns in series, also achieves a greater uniformity of gain between a group of various diameter helices designed for various channels but using a common mast diameter such as is usually available only in discrete sizes. The use of an extremely large number of turns per bay requires a very close spacing between mast and helix in terms of wave lengths so that the radiation per turn can be kept low enough to secure full illumination of the bay. As the helix diameter is altered to achieve turn length for other channels in the same channel group, a large percentage change occurs in this spacing, and hence in the illumination of the bay, with a consequent variation of the gain between designs for various channels based on that mast size. The use of a more moderate number of turns per bay results in a greater spacing between helix and mast for proper radiation per turn. Consequently, a given helix diameter change

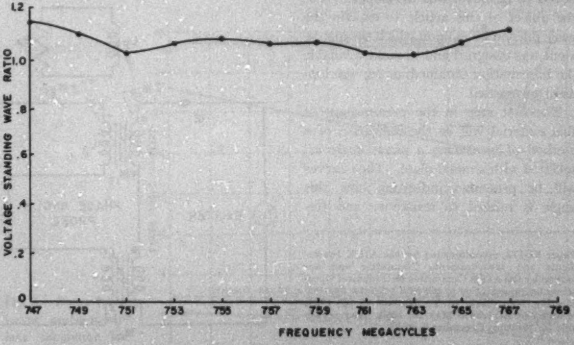


Fig. 4. Plot of voltage standing wave characteristic for antenna operating on channel 61

to accommodate other channels does not result in such a large proportionate change in the spacing to the mast, and hence does not result in as great a change in the illumination of the bay and change in gain of the bay. This means that a greater number of channel bay designs can be accomodated with a given mast size with tolerable variation in gain between specific designs.

In the practical construction of such a multibay array the mast itself may be used as the outer conductor of a coaxial line for feeding the individual bays. Adjustable coupling and compensating reactance is provided at each bay feed point and slug transformers are utilized to secure proper power division between bays and the reduction of reactive storages internal to the mast. Series isolation of the various feed points is provided so that power frequency deicing currents can be passed through the combined helices of a multibay array connected together in series. Lighting wiring for beacons may be passed up through the inside of the inner conductor. Fig. 3 shows such an antenna temporarily supported for photographic purposes by a rigging at its upper end. In actual installation rigging, use is made of the supporting lug visible at the lower end of the second bay from the top.

The whole device operates at a relatively low impedance level throughout with the consequent freedom from the precautions necessary in devices containing the voltage and current concentrations inherent in systems with high reactive energy storages. A plot of the measured voltage standing wave ratio at the drive point of a 5-bay channel 61 antenna is shown in Fig. 4. It therefore comes very close to being a truly travelling wave device making the mode transformation from a coaxial line to the form of cylindrical waves in space without resort to a resonant system at any point. It appears to have sufficient promise even to extend its use down into the very-high-frequency television and frequency-modulated regions where its simple configuration may be adaptable to multiple antenna installations upon a single mast.

No Discussion

Sheet and Plated-Metal Measurements with a Phase-Angle-Type Probe

W. A. YATES
NONMEMBER AIEE

J. L. QUEEN
NONMEMBER AIEE

Synopsis: The basic operating principles and design data for an instrument to measure metal thickness or plating depth for nonferrous materials are described. The mode of operation, involving the measurement of a phase angle, permits the use of a noncontacting probe suitable for the continuous inspection of samples in motion.

IN THE course of the design of a plating-thickness indicator for a particular type of sample, an operating principle believed to be novel was developed. It is the object of this article to explain the basic phenomenon upon which the instrument was designed and to make available the information obtained as the development progressed.

The first step in the presentation of this material will be the indication of a method of measuring a phase angle associated with a metal plate. Then curves will be presented indicating how this angle is related to resistivity and frequency for metal samples of infinite area and thickness. Next the case of thin metal plates will be considered, and then the case of one metal plated upon another.

Phase-Angle Probe

The basic assumption upon which the plating-thickness indicator was designed is that, as far as influence upon a mutual-inductance transducer is concerned, a nonmagnetic metal plate may be replaced by an equivalent coil, L_p, shunted by an equivalent resistance R_p. The equivalent plate impedance $R_p+j\omega L_p$ is primarily a function of the transducer coefficients, the excitation frequency, and the plate dimensions and resistivity. For a given transducer and frequency, the equivalent impedance of a plate of sufficient area will be a function of its thickness and resistivity. For constant resistivity the impedance becomes a function of thickness only, so that thickness of a known plate may be obtained from measurement of the plate impedance amplitude or phase angle alone, assuming they are monotonic functions of thickness.

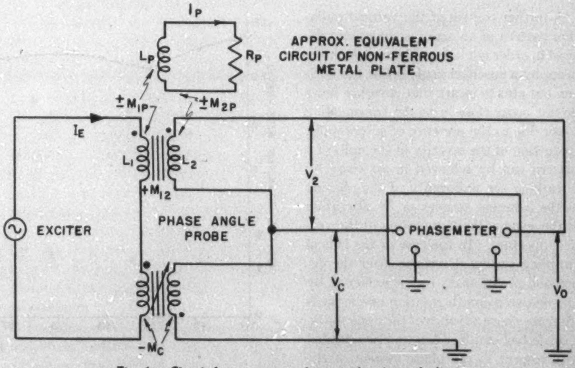
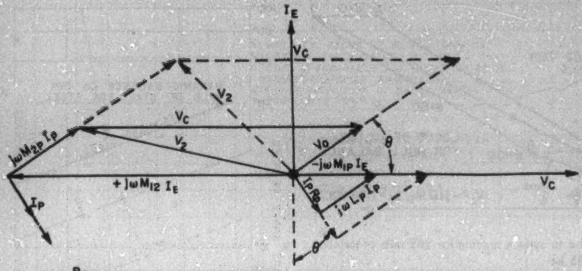


Fig. 1. Circuit for measuring phase angle of metal plate

Paper 54-173, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted May 29, 1953; made available for printing December 9, 1953.

W. A. YATES and J. L. QUEEN are with the National Bureau of Standards, Washington, D. C.



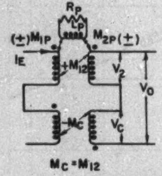


Fig. 2. Phasor diagram for phase-angle probe near coil L_p and resistor R_p in shunt. Dotted lines are for tight coupling between transducer and coil; solid lines are for loose coupling

$$V_c = -j\omega M_c I_E$$
$$V_2 = j\omega M_{12} I_E + j\omega M_{2p} I_p$$
$$V_0 = V_c + V_2 = j\omega M_{2p} I_p$$

$$I_p = \frac{-j\omega M_{1p} I_E}{R_p + j\omega L_p}$$

$$\theta = \text{arc tan}\ \frac{R_p}{WL_p}$$

A phase-angle measurement has one distinct advantage, insensitivity to transducer-to-metal spacing. This advantage is realized by use of the phase-angle probe, shown in Fig. 1, which is a combination of a mutual-inductance transducer and a compensating transformer so arranged as to give two output voltages, V_0 and V_c, whose phase difference is the complement of the plate impedance phase angle. With the plate removed from the area of influence of the transducer, M_c is made equal to M_{12} by adjusting it for a null at V_0, so that with the plate in proximity

$$V_0 = V_c + V_2 = -j\omega M_c I_E + (+j\omega M_{12} I_E + j\omega M_{2p} I_p) = j\omega M_{2p} I_p$$

Since

$$I_p Z_p + j\omega(+M_{1p}) I_E = 0$$

$$I_p = \frac{j\omega M_{1p} I_E}{Z_p}$$

$$V_0 = \left[\frac{\omega M_{1p} M_{2p}}{M_c(R_p^2 + \omega^2 L_p^2)}\right] V_c(\omega L_p + jR_p)$$

or using

$$M_{1p} = K_1\sqrt{L_1 L_p},\ M_{2p} = K_2\sqrt{L_2 L_p},$$
$$M_c = M_{12} = K\sqrt{L_1 L_2}$$

$$V_0 = V_c\left[\frac{K_1 K_2}{K}\cos\theta\right] \angle\theta$$

where

$$\theta = \text{arc tan}\ \frac{R_p}{\omega L_p} = \frac{\pi}{2} - \text{arc tan}\ \frac{\omega L_p}{R_p}$$

That is, the phase difference of the two output voltages is equal to the complement of the plate-impedance angle and independent of the probe-to-plate coupling and primary excitation. Fig. 2 is a phasor diagram of the voltages and currents of the phase-angle probe.

Probe Balancing

Phase-angle measurements on nonmagnetic metal plates with the circuit of Fig. 1 indicated that replacement of a plate by an equivalent coil and resistor was valid only to a first-order approximation, and that the probe-to-plate spacing had a second-order effect in the determination of the phase angle. However, because of the manner in which the phase angle varies with probe-to-surface separation, a major part of the variation can be compensated for by an appropriate unbalance between M_c and M_{12}. The necessary amount of unbalance may best be determined by changing the probe-to-surface spacing while adjusting M_c to give a minimum deviation in phase-angle reading. This method of probe alignment is referred to as dynamic balancing.

Fig. 3 shows curves of phase angle versus probe-to-surface spacing for stainless steel and copper, using both null and dynamic balancing, and indicates the improvement obtained by dynamic balancing in each case.

Fig. 4 is a phasor diagram of $V_0\underline{/\theta_d}$ for a stainless-steel sample showing why the improvement is obtained. From this diagram it can be seen that although the magnitude and phase angle of V_0 decrease as the probe-to-surface spacing is increased, the locus of the end point of the phasor is essentially a straight line AB. By extending AB until it intersects the V_c phasor, a quantity $OC\underline{/0°}$ is obtained such that $V_0 - OC$ has a phase angle θ which is independent of probe-to-surface spacing. The dynamic-balancing process is a practical way of obtaining this condition. From null balance it is necessary to reduce M_c by the ratio of OC to V_c to obtain a dynamic balance. For the 262-mil stainless-steel sample of Fig. 4 this reduction amounted to 4 per cent. The amount of compensation is to some extent dependent on the phase-angle magnitude so that a compromise must be made when readings are desired on metals covering wide ranges of resistivity and thickness.

Empirical Equation for Phase Angle of Metals in Limited Range of Resistivity and Thickness

To determine the manner in which the phase angle depends on frequency, resistivity, and thickness, a number of measurements were made on samples ranging from copper to mercury (1.7 to 96 mi-

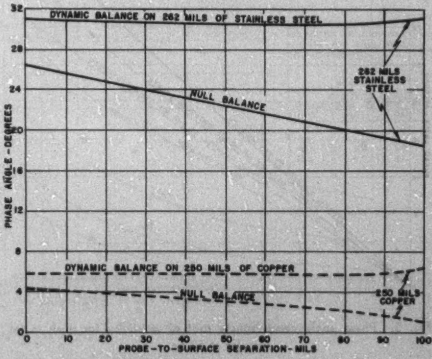


Fig. 3. Phase angle versus probe-to-surface separation for copper and stainless steel at 25 kc

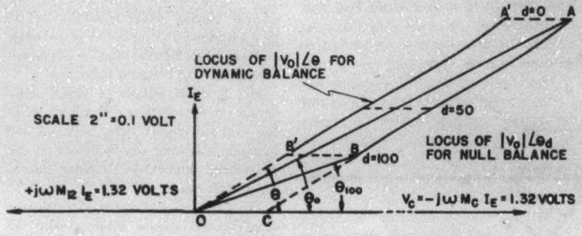

A'
d=0
A
LOCUS OF $|V_0|\angle\theta$ FOR DYNAMIC BALANCE
I_E
SCALE 2" = 0.1 VOLT
d=50
LOCUS OF $|V_0|\angle\theta_d$ FOR NULL BALANCE
B'
B
d=100
$+j\omega M_{12} I_E = 1.32$ VOLTS
θ
θ_0
θ_{100}
$V_C = -j\omega M_C I_E = 1.32$ VOLTS
O
C

Fig. 4. Compensation of residual effects of probe to surface spacing for 262 mils of stainless steel at 25 kc

d = probe to surface separation in mils

OC = B'B = A'A = 0.05 volts = amount V_o must be reduced from null balance condition to produce a dynamic balance

crohm-centimeters resistivity) using the circuit of Fig. 1. The exciter consisted of an *HP 200-C* oscillator driving a McIntosh *50-W-2* amplifier. The phase meter was one specifically designed for the plating-thickness indicator, and was calibrated for the desired frequency and input amplitude ranges. The readings were obtained with an accuracy of better than 5 per cent. The mutual-inductance transducer was wound on a 1/4-inch diameter Stackpole core of grade S-49 powdered iron. The primary was 60 turns of no. 30 wire; the secondary 275 turns of no. 38 wire. The length of winding was 0.2 inch, and was located at the end of the core. The compensating transformer was similarly wound but with a 250-turn secondary. Variation of M_c was accomplished by adjusting a second piece of powdered iron in proximity to the compensating transformer. All measurements were made with the transducer in physical contact with the samples.

Fig. 5 is an experimental curve of phase angle versus square root of resistivity at constant frequency. Fig. 6 is phase angle versus frequency at constant resistivity. In both cases the samples were of sufficient thickness that the current in the plate was essentially zero at the back surface. For dynamic balancing on stainless steel, it is found from Fig. 5 that the phase angle of a thick metal plate at 50 kc is

$$\theta = 2.77\sqrt{\rho}$$

where

θ = phase angle in degrees
ρ = resistivity of the sample in microhm-centimeters (cm)

Using this equation and the curves of Fig. 6, the phase angle of a thick metal plate is

$$\theta = 2.77\sqrt{\rho}\left[\frac{50{,}000}{f}\right]^{0.364}$$

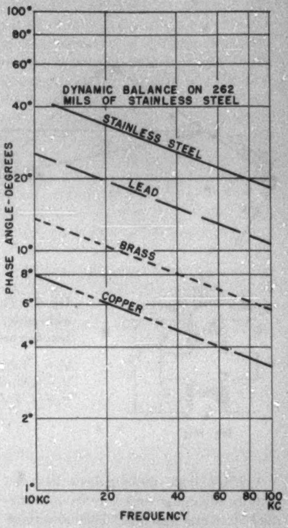


Fig. 6. Phase angle versus frequency for thick nonmagnetic metals

$$\theta = 0.313\delta\left[\frac{f}{50{,}000}\right]^{0.136}$$

where θ and ρ are defined as in the foregoing and

f = frequency in cycles per second
$\delta = 1{,}980\sqrt{\rho/f}$ = skin depth in mils assuming the magnetic permeability of the plate equals unity.

Fig. 7 contains curves of phase angle

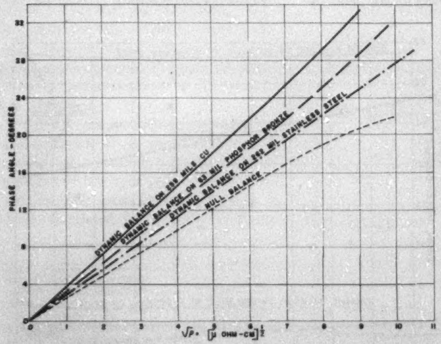


Fig. 5. Phase angle versus square root of resistivity for thick metal samples at 50 kc

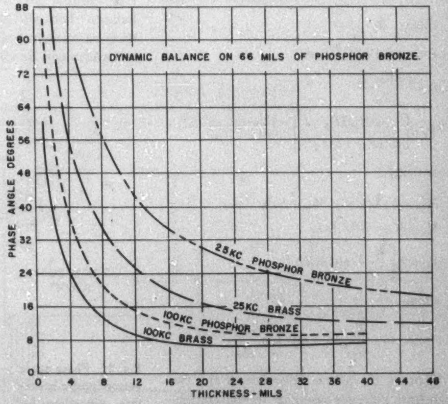


Fig. 7. Phase angle versus thickness for brass and phosphor bronze at 25 and 100 kc

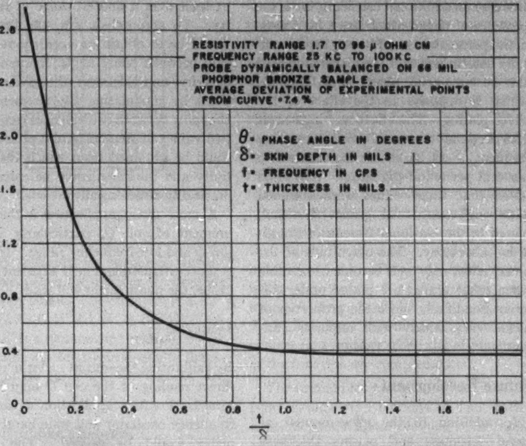


Fig. 8. $\frac{\theta}{\delta}\left(\frac{50,000}{f}\right)^{0.137}$ versus $\frac{t}{\delta}$

versus thickness for brass and phosphor bronze at 25 and 100 kc. Similar curves were obtained for copper, stainless steel, lead, mercury, and a lead zinc alloy. By first plotting θ versus t/δ and then applying a frequency term, the curve of

$$\frac{\theta}{\delta}\left[\frac{50,000}{f}\right]^{0.137}$$

versus t/δ shown in Fig. 8 was developed. This single curve represents all these metals having resistivities between 1.72 and 96 microhm-cm over a frequency range of from 25 to 100 kc. The curve of Fig. 8 was plotted on log-log paper as shown in Fig. 9 and the equation of the straight portion evaluated as

$$\frac{\theta}{\delta}\left[\frac{50,000}{f}\right]^{0.137}=0.378\left[\frac{t}{\delta}\right]^{0.8}$$

or

$$\theta=7.35\times10^{4}\frac{\rho^{0.9}}{t^{0.8}f^{0.763}}$$

for

$$0.15<\frac{t}{\delta}<0.85$$

where θ is expressed in degrees, ρ in microhm-cm, t in mils, and frequency in cycles per second. It should be emphasized that this equation is valid only for the designated probe dynamically balanced on 66 mils of phosphor bronze, although similar equations can undoubtedly be derived for other probe configurations.

Metal-Thickness Measurement

Choice of operating conditions for a metal-thickness indicator to cover a specific range of thickness of one material is governed by many factors. Although sufficient information is not known about the effects of probe parameters and balance, as well as resistivity, frequency, and thickness, on phase angle to allow generalization of optimum operating conditions under all circumstances, certain conclusions may be drawn from the limited amount of data obtained with one specific probe. For a given thickness of material, very little sensitivity is realized until the frequency is reduced to a point where the calculated skin depth of the material is greater than the thickness. Further reduction of frequency results in an increase in absolute sensitivity at approximately constant percentage sensitivity but a decrease in V_c and the ratio of V_0 to V_c. How far the frequency can be decreased with advantage depends upon the maximum amount of undistorted exciter current available, the minimum amplitude required for phase-meter operation, the precision with which the probe may be balanced, and the probe parameters. As metal thickness and resistivity increase, a larger probe diameter is desirable. For measurement of thickness of one metal only, the output meter may be calibrated directly in thickness. For a general - purpose instrument, multiple meter graduations or calibration curves could be used. It may be possible, however, to avoid this expedient and have a single scale reading thickness directly, but only at the expense of more complex circuitry.

Plating-Thickness Measurement

Choice of operating conditions for a plating-thickness indicator is strongly influenced by particular circumstances, since the plating and base metal both con-

Fig. 9. $\frac{\theta}{\delta}\left(\frac{50,000}{f}\right)^{0.137}$ versus $\frac{t}{\delta}$ for nonmagnetic metals

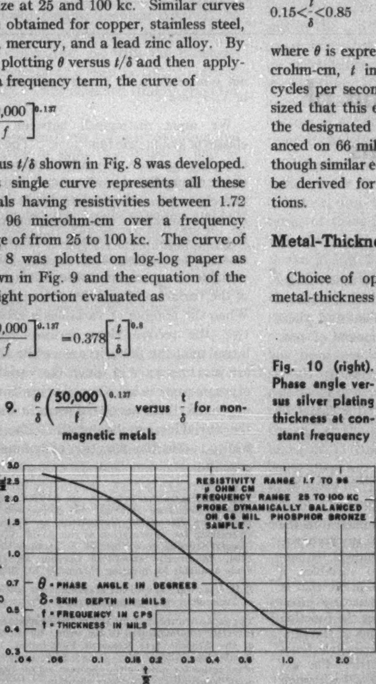
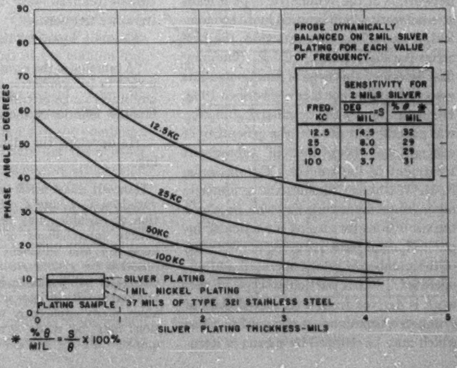


Fig. 10 (right). Phase angle versus silver plating thickness at constant frequency



FREQ. KC	SENSITIVITY FOR 2 MILS SILVER	
	$\frac{DEG}{MIL}=S$	$\frac{\%\theta}{MIL}$ *
12.5	14.5	32
25	8.0	29
50	5.0	29
100	3.7	31

tribute to the phase angle. For a thin plating of low resistivity on a very thick base metal of high resistivity, the problem is essentially the same as for metal thickness alone. As the resistivities approach the same value, difficulties are encountered, but in present plating applications, this is not the usual case. If the thickness of the base material varies, the frequency must be raised until the transducer field does not penetrate the thinnest portion of the base material, whereas if the base thickness is constant a higher absolute sensitivity may be gained by lowering the frequency.

Fig. 10 demonstrates the variation of phase angle with silver-plating thickness on 1 mil of nickel plated on 37 mils of stainless steel. The curves indicate the values of phase angle of the base material at zero silver plating and approach the phase angle of silver alone as the silver thickness increases. As can be seen, the absolute sensitivity is increasing as the frequency is decreased while the per-cent sensitivity is remaining constant, so that the characteristics of the phase meter are of prime importance in determining the operating frequency. In this case an additional consideration is the presence of nickel which has a high magnetic permeability. It was experimentally determined that the nickel increased the base-material phase angle slightly, which is opposite to the effect of the silver plating. The effect was highest at 12.5 kc and decreased with increasing frequency. At 25 kc the increase was 5 per cent per mil of the nickel. It is to be noted that these values were obtained with one commercial grade of nickel, plated in the National Bureau of Standards Laboratory. The magnitude of the nickel effect may well become intolerable when other grades are plated under different conditions, since the properties of nickel vary widely with impurity content.

Future Development

In addition to the measurement of metal thickness and plating thickness where the materials are known, and calibration curves have previously been obtained, there is a possibility of developing an instrument which would measure thickness and resistivity simultaneously, at least over a limited range of resistivity. To accomplish this, the first step would be to obtain an accurate relationship between phase angle, frequency, resistivity, and thickness of a given probe. Then by modulating the frequency, or chopping between two values of frequency, two values of frequency and phase angle could be obtained from which the resistivity and thickness could be calculated, preferably electronically.

A more direct application is the measurement of coil Q, particularly in the low-Q and low-frequency range. If a coil is short-circuited and held adjacent to the probe, the phase meter will read

$$\theta = \text{arc tan}\ \frac{R}{\omega L}$$

By calibrating the output meter in cot θ, a direct reading of the coil Q may be obtained. For this application a variable-frequency oscillator and wide-band phase meter would be needed. The probe would be null-balanced.

No Discussion

The Recognition and Identification of Symmetric Switching Functions

S. H. CALDWELL
ASSOCIATE MEMBER AIEE

In the design of switching circuits based on relay-contact networks, it is known that if the function which describes the network is of the symmetric type it leads to a network which is much more economical of switching elements than the best series-parallel circuit. C. E. Shannon[1] first described this type of function and its associated relay contact network in 1938. His definition of the symmetric function was general, but most of the generality of concept has been lost in subsequent treatment of the subject. For example, S. H. Washburn[2] defines the symmetric relay-contact network as "one in which the conditions for closing a particular input-to-output path are given ... in terms of the number of relays operated and unoperated." This particular type of network is derived from a restricted class of symmetric functions—a class of functions which may be realized by means of iterative networks.[3] There is a much larger class of symmetric functions contained within Shannon's definition, and these can be realized only by means of noniterative networks.

Shannon's general definition is:

> "A function of the n variables $X_1, X_2 \ldots X_n$ is said to be symmetric in these variables if any interchange of the variables leaves the function identically the same."

He develops the idea further by stating:

> "Thus $XY+XZ+YZ$ is symmetric in the variables X, Y, and Z. Since any permutation of variables may be obtained by successive interchanges of two variables, a necessary and sufficient condition that a function be symmetric is that any interchange of two variables leaves the function unaltered. By proper selection of the variables, many apparently unsymmetric functions may be made symmetric. For example, $XY'Z+X'YZ+X'Y'Z'$ although not symmetric in X, Y, and Z is symmetric in X, Y, and Z'."

We must distinguish between two classes of symmetric functions. The first class has variables of symmetry which are either all unprimed or all primed, and it is this class of functions which leads to the iterative networks which are characteristic of "m out of n" circuits. The second class will be designated as the mixed variable class, because some, but not all, of the variables of symmetry are primed. When the number of variables n exceeds two, the second class becomes much larger than the first, because while there are just two ways in which the variables of symmetry can be either all primed or all unprimed, there are 2^n-2 ways in which the variables can be mixed. The difficulty is that the detection of symmetric

---

Paper 54-24, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 16, 1953; made available for printing November 13, 1953.

S. H. Caldwell is with the Massachusetts Institute of Technology, Cambridge, Mass.

The expansion theorem for symmetric functions was privately communicated to the writer by R. Grea of the Graphic Arts Research Foundation, Inc., Cambridge, Mass. Mr. Grea used a method of induction for proof of the theorem, rather than a proof based on the a-number theorem.

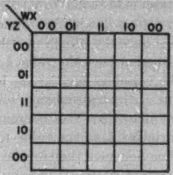


Fig. 1. Extended form of Karnaugh map

functions in mixed variables becomes exceedingly difficult by algebraic procedures, even when the number of variables is relatively small. For example, $WXYZ' + WX'YZ + W'XY'Z + WX'Y'Z' + W'X'YZ'$ is a symmetric function, but it is no simple matter algebraically to determine that its variables of symmetry are W, X', Y, Z'.

The *a*-Number Theorem

The proposed method for recognizing both classes of symmetric circuits starts with the *a*-number theorem given by Shannon.[1] Originally, the theorem was stated in terms of the hindrance concept (0 = closed circuit, 1 = open circuit); as stated in the following, it is modified for more convenient use with the transmission concept (0 = open circuit, 1 = closed circuit). Thus modified (italic type), the theorem is:

> "A necessary and sufficient condition that a function be symmetric is that it may be specified by stating a set of numbers a_1, $a_2 \ldots a_k$ such that if exactly $a_j (j = 1, 2, 3 \ldots k)$ of the variables *have the value one*, then the function *has the value one*, and not otherwise. . . The set of numbers a_1, $a_2 \ldots a_k$ may be any set of numbers selected from the numbers 0 to n, inclusive, where n is the number of variables in the symmetric function. For convenience, they will be called the *a*-numbers of the function."

The symmetric function $XY+XZ+YZ$ has the *a*-numbers 2 and 3, since the function has the value 1 if either two or three of the variables have the value 1, but not if one or if none have the value 1. Symbolically, we express this function as $S_{2,3}(X,Y,Z)$. The function $XY'Z + X'YZ + X'Y'Z'$ has the mixed variables of symmetry X,Y,Z' and the *a*-number 1, so it is written $S_1(X,Y,Z')$. It may also be verified that the illustrative function of four variables given above is $S_{0,3}(W,X',Y,Z')$.

Symmetric Functions of Four Variables

By means of an expansion theorem, the problem of recognizing symmetric functions of more than four variables will be reduced to a problem in exactly four variables. Consequently, we are concerned here with some of the special properties of symmetric functions of four variables. Every symmetric function of four variables can be written in two ways. The function $S_1(W,X,Y,Z)$ is identically equal to the function $S_3(W',X',Y',Z')$. The rule is that a symmetric function of any four variables is equal to the symmetric function of the corresponding complementary variables when each *a*-number a_j of the original function is replaced by the *a*-number $(4-a_j)$. This rule holds even when the variables of symmetry are mixed. Thus $S_{1,4}(W',X,Y,Z') = S_{0,3}(W, X',Y',Z)$. For the *a*-number 2, there is no change in *a*-number when the arguments of the function are complemented.

The method for recognizing and identifying symmetric functions of four variables depends upon the use of a map which is an extension of the chart described by E. W. Veitch[4] and later refined by M. Karnaugh.[5] If Karnaugh's map for four variables is extended by the addition of a 00 row and a 00 column, as shown in Fig. 1, and entries on the map are duplicated in the left and right columns and in the top and bottom rows, it becomes somewhat easier to observe the adjacencies between the terms of a switching function. It is no longer necessary to keep in mind that entries at the top and bottom of a column are adjacent, as are those at the left and right ends of a row.

The extended map of Fig. 1 is quite sufficient for the detection of symmetric functions, but the relations among the terms of these functions become clearer when the map is extended still further, as shown in Fig. 2. The same cyclic order of number pairs is repeated for the additional rows and columns, and again the map is terminated in a 00 row and a 00 column. In each cell of the map a number has been placed. This number corresponds to the total number of 1's in the combined designators of the column and row which intersect at that cell. Thus, the cell at the intersection of column 01 and row 11 contains the number 3 because there are three 1's in the combined designator 0111. There are three other sets of combined designators which lead to the number 3 in their cells; namely, 1011, 1101, and 1110. We observe now that the number 3 always occurs in groups of four, and that each group consists of a small diagonal square array of entries; these arrays are emphasized by the diagonal lines drawn in Fig. 2. If a given switching function exhibits transmission values of 1 in all four of these cells, and in no others, it satisfies the *a*-number theorem and the function may be identified as $S_3(W,X,Y,Z)$. This is evident from the fact that the array of 3-cells represents all the ways in which exactly three of the variables W,X,Y and Z can have the value 1.

In similar manner the arrays of 1-cells and of 2-cells shown in Fig. 2 show the location of transmission points (points for which the function has the value 1) of symmetric functions of W,X,Y, and Z, having *a*-numbers of 1 and 2, respectively. The *a*-number 2 is readily distinguished because it occurs in a large diagonal-square array of entries, rather than the small diagonal-square arrays which indicate *a*-numbers of 1 or 3. Since there is only one term in a function of *a*-number 0, and one term in a function of *a*-number 4, these transmission points are isolated from each other, but the arrays for *a*-numbers 1, 2, and 3 are located symmetrically about them.

Functions with Mixed Variables of Symmetry

We now examine what happens when we map a function which is symmetric in W,X,Y,Z, and then map the same function when it has mixed variables of symmetry. Consider the functions

$$S_{2,3}(W,X,Y,Z) = WXY'Z' + WX'YZ' + WX'Y'Z + W'XYZ' + W'XY'Z + W'X'YZ + WXYZ' + WXY'Z + WX'YZ + W'XYZ \quad (1)$$

$$S_{2,3}(W,X',Y',Z) = WX'YZ' + WXY'Z' + WXYZ + W'X'Y'Z' + W'X'YZ + W'XY'Z + WX'Y'Z' + WX'YZ + WXY'Z + W'X'Y'Z \quad (2)$$

Function 1 is mapped in Fig. 3(A) and it can be observed that the transmission points form a pattern of large and small diagonal-square arrays. From a

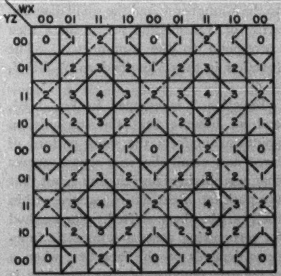
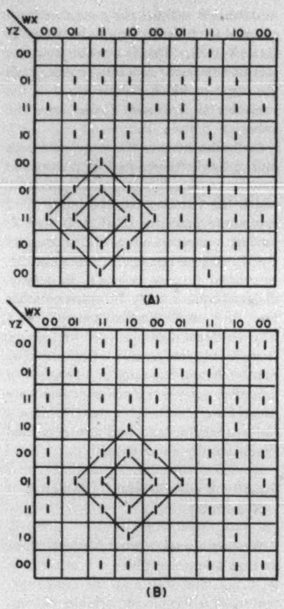


Fig. 2. Extended map for detection of symmetric functions



(A)

(B)

Fig. 3. Symmetric function map showing shift in location of pattern when variables of symmetry are changed

comparison with Fig. 2, it is clear that the large array definitely corresponds to the a-number 2, but the small array may indicate an a-number of either 1 or 3. Both arrays are symmetrically located about the cell designated 1111, and since this cell corresponds to an a-number of either 0 or 4, the variables represented by the cell must be one of the two possible sets of variables of symmetry. If the 1111 designation is interpreted to stand for the term $WXYZ$, the central cell corresponds to the a-number 4, and the small diagonal array of transmission points must then correspond to the a-number 3. The function is thus identified as $S_{2,3}(W,X,Y,Z)$.

Function 2 is mapped in Fig. 3(B), and we observe that the transmission points again form the same pattern of large and small diagonal-square arrays that was obtained for function 1, but that the position of the entire pattern within the chart has been shifted. The method for recognizing general symmetric functions depends on the fact that when a symmetric function of four variables is mapped on an extended chart, using Karnaugh's cyclic ordering of the row and column designators, the transmission points form large and/or small diagonal-square arrays for a-numbers other than 0 or 4; and on the fact that when the variables of symmetry of the function are changed, the entire pattern shifts within the map, but without change in shape. The a-numbers and the variables of symmetry are determined from the row and column designators of the cell, about which one or more patterns are symmetrically located.

In Fig. 3(B), the large diagonal-square array again indicates an a-number of 2, and the small diagonal square may correspond to an a-number of either 1 or 3. Both patterns are symmetrically located about the 1001-cell, which can indicate the variables of symmetry W,X',Y',Z. The a-number of the small diagonal square is determined by testing one of its points against the selected variables of symmetry. Consider the point 1000 in the small diagonal square, which in map co-ordinates represents the term $WX'Y'Z'$. When we compare this term with the variables of symmetry W,X',Y',Z, we see that the test term has a transmission value of 1 when the three variables W,X', and Y' have the value 1, and Z has the value 0. Hence, by Shannon's a-number theorem, the a-number represented by the small diagonal square array is 3, and the function is completely identified as $S_{2,3}(W,X',Y',Z)$. Of course, it could just as well have been identified as $S_{1,2}(W',X,Y,Z')$, which is identically the same function.

Symmetric Functions of More Than Four Variables

When the number of variables of symmetry exceeds four, the direct recognition of symmetric functions by chart patterns becomes quite difficult. An expansion theorem can be derived, however, which enables us to reduce the number of variables to be mapped. Shannon[1] gives a general expansion theorem in the form

$$f(X_1, X_2 \ldots X_n) = X_1[f(1, X_2 \ldots X_n)] + X_1'[f(0, X_2 \ldots X_n)]$$

This theorem may be used to expand a symmetric function of n variables which has the single a-number p. The expansion can be performed about any of the variables, but for simplicity we expand about X_n, and write

$$S_p(X_1, X_2 \ldots X_n) = X_n[S_p(X_1 \ldots X_{n-1}, 1)] + X_n'[S_p(X_1 \ldots X_{n-1}, 0)]$$

The residual function, $S_p(X_1 \ldots X_{n-1}, 1)$, is a symmetric function of a-number equal to p, in which the value of one variable has been made equal to 1. Since the a-number defines the number of variables which must have the value 1 in order that the transmission have the value 1, if any one of the variables such as X_n has the fixed value 1, the transmission will have the value 1 if any $(p-1)$ of the remaining $(n-1)$ variables have the value 1. This residual function must therefore be a symmetric function of $(n-1)$ variables with an a-number $(p-1)$, and the first term of the expansion is written $S_{p-1}(X_1 \ldots X_{n-1})$. We must note, however, that if $p=0$, this a-number vanishes in the $S_{p-1}(X_1 \ldots X_{n-1})$ residual function. This can be seen from the fact that X_n is not a factor of $S_0(X_1 \ldots X_n) = X_1' X_2' \ldots X_n'$.

The other residual function, $S_p(X_1 \ldots X_{n-1}, 0)$, contains only $(n-1)$ variables, because X_n has been made zero. Since none of the variables has been given the fixed value 1, the a-number remains unchanged and this residual function is hence a symmetric function of $(n-1)$ variables with the a-number p, and is written $S_p(X_1 \ldots X_{n-1})$. Here we note that if $p=n$, this a-number must vanish in the $S_p(X_1 \ldots X_{n-1})$ residual function. This can be seen from the fact that X_n' is not a factor of $S_n(X_1 \ldots X_n) = X_1 X_2 \ldots X_n$.

Combining the two terms, we write the expansion theorem for a symmetric function as

$$S_p(X_1 \ldots X_n) = X_n S_{p-1}(X_1 \ldots X_{n-1}) + X_n' S_p(X_1 \ldots X_{n-1}) \quad (3)$$

If a symmetric function has two a-numbers, the expansion theorem takes the form

$$S_{p,q}(X_1 \ldots X_n) = X_n S_{p-1,q-1}(X_1 \ldots X_{n-1}) + X_n' S_{p,q}(X_1 \ldots X_{n-1}) \quad (4)$$

This form of theorem 4 follows directly from the relation

$$S_{p,q}(X_1 \ldots X_n) = S_p(X_1 \ldots X_n) + S_q(X_1 \ldots X_n)$$

The expansion theorems for functions with more than two a-numbers can readily be derived by analogy with theorem 4.

Since the expansion theorems can be applied by expanding successively about any number of variables, we can always obtain a set of residual functions of four variables. If these residual functions are themselves symmetric with respect to a common set of variables, and have the a-number relations required by the expansion theorem, then the original function must be symmetric. Both the a-numbers and the variables of symmetry of the original function can be determined from the chart patterns of the residual functions.

As an example, a function of five variables will be mapped to identify its variables of symmetry and to determine its a-numbers. The function is

$$T = A'BCDE + ABCD'E + ABCDE' + AB'C'DE + A'B'CD'E' + A'B'CDE' + A'BC'D'E + A'BC'DE' + AB'CD'E' + ABC'D'E' + A'B'C'D'E'$$

If the function is symmetric, the outcome of the expansion process is independent of the variable about which the function is expanded. For convenience, expand about the variable E to obtain residual functions of the remaining four variables. Since the given function is in the standard sum form, the expansion can be obtained by simple factoring, as follows

$$T = E(A'BCD + ABCD' + AB'C'D + A'B'CD' + A'BC'D') + E'(ABCD + A'B'CD + A'BC'D + AB'CD' + ABC'D' + A'B'C'D')$$

The residual functions of A,B,C,D are mapped in Fig. 4. In the upper chart, the terms associated with E as a factor present a pattern of small diagonal squares and isolated transmission points. This pattern identifies a symmetric function in which the diagonal squares represent an a-number of either 1 or 3, and the isolated points represent an a-number of either 0 or 4. In the lower chart, the terms associated with E' as a factor present the large diagonal-square pattern which indicates a symmetric function of a-number 2.

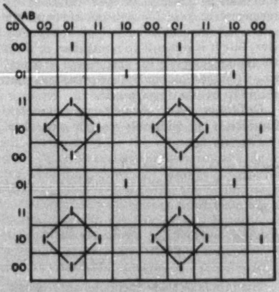
Although the residual functions are now identified as symmetric functions, they do not necessarily satisfy the requirements of the expansion theorem, equation 4. One apparent contradiction is the fact that two a-numbers are indicated by the upper chart, and only one a-number appears in the lower chart. This combination can still satisfy the expansion theorem, however, if one of the original a-numbers is 0. If $p=0$, in the expansion, the a-number $(p-1)$ must vanish. If we assume the isolated points in the upper chart indicate an a-number of 0, then the pattern indicates that the small diagonal squares must represent an a-number 3. For this a-number, $(p-1) = 3-1$, and this is consistent with the a-number 2 shown by the lower chart.

The pattern of the upper chart now indicates that four of the variables of symmetry must be A',B,C,D', because the diagonal-square pattern surrounds the point 0110. The point 0110 is also one of the two central points in the pattern of the lower chart, so both residual functions can be written in terms of the same variables of symmetry.

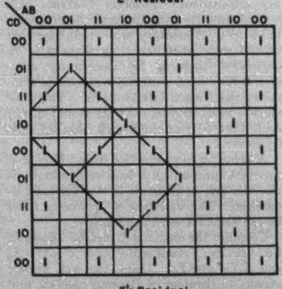
We can now write the function in the following form

$$T = ES_{0,3}(A',B,C,D') + E'S_2(A',B,C,D')$$

Comparing this expression with the expansion theorem, equation 4, we find that the theorem is satisfied if the fifth variable of symmetry is E' instead of E. Hence the original function is identified as $S_{0,3}(A',B,C,D',E')$.



E - Residual



E'- Residual

Fig. 4. Residual functions derived from expansion of a function of five variables

An Example in Six Variables

Consider the expansion of a symmetric function of six variables, when expanded about two of its variables

$$S_p(u,v,w,x,y,z) = uvS_{p-2}(w,x,y,z) + uv'S_{p-1}(w,x,y,z) + u'vS_{p-1}(w,x,y,z) + u'v'S_p(w,x,y,z) \quad (5)$$

The variables u,v,w,x,y,z are generic and may actually be mixed. If there is more than one a-number, an equation similar to equation 4 can be written.

Let us designate the four residual functions as the -2 residual, -1 residuals, and the -0 residual, referring to the reduction of the a-numbers in the expansion process. The a-numbers 0 and 1 will vanish in the -2 residual, 0 and 6 will vanish in the -1 residual, and 5 and 6 will vanish in the -0 residual.

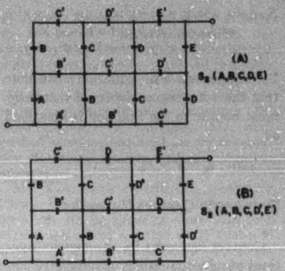


Fig. 5. Effect on symmetric network of a primed variable of symmetry

We note from equation 5 that all the residuals are functions of the same variables, and that two of the residuals must have the same a-numbers. These observations enable us, from maps of the four residuals, to identify any symmetric function of six variables.

Suppose we have mapped the residuals of a 6-variable function after expansion about the variables u and v, and have identified the following

uv residual $= S_{0,2,4}(w',x',y,z)$
uv' residual $= S_{1,3}(w',x',y,z)$
$u'v$ residual $= S_{1,3,4}(w',x',y,z)$
$u'v'$ residual $= S_{0,2,4}(w',x',y,z)$

These residuals are all functions of the same variables, w',x',y,z, as required by equation 5. It remains to be seen whether the a-numbers form a consistent set.

We do not know how many a-numbers there are in the 6-variable function, so we denote the possible a-numbers by $p,q,r,s, \ldots$ The uv-residual and the $u'v'$ residual both have the a-numbers 0,2,4. Hence, these must be the -1 residuals, and we can therefore identify $p,q,r = 1,3,5$ as some of the required a-numbers.

This result suggests that the a-numbers 1,3 of the uv' residual may represent $(q-2)$, $(r-2)$; $(p-1)$ vanishes. But in that case, the $u'v$ residual is the -0 residual, and its a-numbers are 1,3,4—a result which is inconsistent, because it would require a 3 in the a-numbers of the two residuals which are known to be the -1 residuals.

The alternative hypothesis is that the $u'v$ residual is the -2 residual. Hence its a-numbers of 1,3,4 become the a-numbers 3,5,6 of the 6-variable function. Hence, we have $p,q,r,s = 1,3,5,6$, and this set is consistent for all residuals.

If the terms are now grouped as in equation 5 we write

$$S_{1,3,5,6}(u,v,w,x,y,z)=u'vS_{1,3,4}(w',x',y,z)+uvS_{0,2,4}(w',x',y,z)+u'v'S_{0,2,4}\times(w',x',y,z)+uv'S_{1,3}(w',x',y,z)$$

By comparison with equation 5 we note that the other two variables of symmetry must be $u'v$, so the function is identified as

$$S_{1,3,5,6}(u',v,w',x',y,z)$$

Symmetric Contact Networks

Contact networks for symmetric functions of mixed variables of symmetry can readily be built from the basic symmetric network. The process is illustrated in Fig. 5, where part (A) shows the network for $S_2(A,B,C,D,E)$, and part (B) shows the network for $S_2(A,B,C,D',E)$. As a practical matter, the required network is drawn by first assuming that all variables of symmetry are unprimed. After the network is designed for that set of variables, the contact designators for those variables of symmetry which are primed are replaced by their complements.

It is obvious that once the symmetric function is identified, the problem of designing the network is trivial. There is, however, a fundamental change in the nature of the resulting network. When the variables of symmetry are all unprimed or all primed, the network can be designed by iterative methods, and we can hence consider it to be composed of a chain of typical cells.[3] This is no longer true if the variables of symmetry are mixed, and this class of functions hence leads to noniterative symmetric networks.

Conclusion

By the use of an extension of the Karnaugh map, together with the a-number theorem and an expansion theorem for symmetric functions, it has been shown that the most general type of symmetric function can be detected and identified. In their development of the map method for treating switching functions, both Veitch and Karnaugh considered only horizontal and vertical arrays of transmission points. This paper shows that diagonal arrays are associated with symmetric functions, and examines the rules for interpreting such arrays. The importance of a recognition and identification method lies in the fact that there are many more symmetric functions in mixed variables than there are of the familiar "m out of n" type.

References

1. A Symbolic Analysis of Relay and Switching Circuits, **Claude E. Shannon.** *AIEE Transactions*, vol. 57, 1938, pp. 713–22

2. Relay "Trees" and Symmetric Circuits, **S. H. Washburn.** *AIEE Transactions*, vol. 68, pt. I, 1949, pp. 582–86.

3. The Design of Switching Circuits (book), **William Keister, A. E. Ritchie, S. H. Washburn.** D. Van Nostrand Company, New York, N. Y., 1951, article 4.10.

4. A Chart Method for Simplifying Truth Functions, **E. W. Veitch.** *Proceedings*, Association for Computing Machinery, Pittsburgh, Pa., May 1952, pp. 128–33.

5. The Map Method for Synthesis of Combinational Logic Circuits, **M. Karnaugh.** *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 593–99.

◆

Discussion

S. H. Washburn (Bell Telephone Laboratories, New York, N. Y.): In the analytical study of switching functions and their corresponding circuit configurations, one of the most interesting classes is that of symmetric functions. Professor Caldwell's paper makes a welcome contribution in presenting certain characteristics of these functions in an integrated form and in suggesting a method for recognizing inherent symmetry in a given function. This method is particularly attractive since it avoids the necessity of assuming an arbitrary reference point in analyzing the function in question.

Professor Caldwell in a number of places emphasizes the distinction between the "m out of n" symmetric class and what he has designated the mixed variable class. It is not clear to me that these two classes are fundamentally different; in fact, it would appear that both in circuit analysis and in circuit synthesis it is more convenient to treat both symmetric situations as a single all-inclusive class.

Consider, for example, the symmetric function $S_1(X,Y,Z')$. The variable Z' may be replaced by a new variable W so that the function becomes $S_1(X,Y,W)$. This represents, in relay contact network terms, a network closed for one out of the three variable relays (X, Y, and W) operated. Generalizing, any symmetric belongs to the "m out of n" class if we are permitted to consider those variables which are primed to be replaced by new unprimed variables. We do this in practice, of course, when we verbally describe the function as being symmetric in X, Y, and Z-prime.

In synthesizing a circuit to correspond to a symmetric function, it is also helpful to avoid the author's distinction. If we replace all primed variables in the S-function by new unprimed variables, the required contact network will be iterative in the new set of variables and may be designed by the standard techniques applicable to iterative networks. Finally, after the resulting network has been drawn, we modify it to correspond to the original set of variables by replacing make-contacts with break-contacts, and vice versa, on those relays designated by the variables which were primed in the original S-function. The wiring interconnecting the contacts remains unchanged.

M. Karnaugh (Bell Telephone Laboratories, Murray Hill, N. J.): The development of methods for efficient logic-circuit synthesis is presently proceeding along two lines. The first of these involves the construction and/or programming of digital computers for this task. The other involves the improvement of rapid desk-top methods. The former is best for complicated problems, while the latter is to be preferred for problems of little or moderate difficulty.

There are some important reasons why most problems will continue to involve only moderate numbers of variables. One reason is the limitations of electronic components, many of which work best in small packages. Another reason is the economies which are realized by serial rather than parallel operation. Finally, maintenance considerations tend to limit circuit complexity. Thus, it is clear that desk-top methods of design will be of major importance in the foreseeable future.

Professor Caldwell's basic contribution is the observation that a symmetric function of as many as four variables is easy to recognize by inspection of its map (see ref. 5 of the paper). This addition to desk-top design techniques is chiefly applicable to relay contact networks at present. However, it is quite possible that other components will be found which can efficiently realize symmetric functions.

The author next extends his method to functions of more than four variables by means of a special expansion theorem for symmetric functions. His approach is effective, but it is difficult to remember, or even to explicitly state, all the rules. I would like to suggest an equivalent procedure that seems to be more systematic.

Definition: An elementary symmetric function is one with only a single a-number.

Theorem 1:

$$S_{\alpha,\beta,\ldots,\nu}(X_1,\ldots X_n)=S_\alpha(X_1,\ldots X_n)+S_\beta(X_1,\ldots X_n)+\ldots+S_\nu(X_1,\ldots X_n)$$

Theorem 2:

$$S_\alpha(X_1,\ldots X_n)=\sum_{i+j=\alpha} S_i(X_1\ldots X_k)\times S_j(X_{k+1},\ldots X_n)$$

The proofs of these theorems are reasonably self-evident. Theorem 1 shows how any symmetric function may be expressed in terms of elementary symmetric functions of the same variables. Theorem 2 is an expansion theorem for elementary symmetric functions, with the variables of symmetry partitioned into two sets. The summation must extend over all values of i, j such that $i+j=\alpha$ and such that $0\leq i\leq k$, $0\leq j\leq n-k$.

A function is symmetric if and only if it can be expressed as a sum of summations like the one in theorem 2, one for each a-number. The variables of symmetry must be the same for each summation.

Consider the author's 6-variable example. For brevity, I shall write $S_\alpha(w',x',y,z)=S_\alpha$ and $S_\alpha(u',v)=T_\alpha$. Using the well known expansion theorem of Boolean algebra, and identifying the residual 4-variable functions as suggested by the author, we get

$$F=u'vS_{1,3,4}+uvS_{0,2,4}+u'v'S_{0,2,4}+uv'S_{1,3}$$

Application of theorem 1, followed by the collection of coefficients, gives

$$F=(uv+u'v')S_0+(uv'+u'v)S_1+(uv+u'v')S_2+(uv'+u'v)S_3+(uv+u'v'+u'v)S_4$$

Now the coefficients must be examined for symmetry. In this case, it is done at a glance, for

$$uv+u'v'=T_1,\ uv'=T_0,\ u'v=T_2$$

Then we have

$$F=T_1S_0+(T_0+T_2)S_1+T_1S_2+(T_0+T_2)S_3+(T_1+T_2)S_4=(T_0S_1+T_1S_0)+(T_0S_3+T_1S_2+T_2S_1)+(T_1S_4+T_2S_3)+T_2S_4=S_{1,3,5,6}(u',v,w',x',y,z)$$

S. H. Caldwell: Mr. Washburn is quite correct in noting that there is only one, all-inclusive class of symmetric functions. The distinction given between "*m* out of *n*" and mixed variable types was based entirely on the relative difficulty of recognizing the existence of a symmetric function if the variables are mixed. A symmetric function of the "*m* out of *n*" type can be detected by the obvious process of counting the number of primed or unprimed variables in each term of the function, when it is written in a fully expanded form.

There must be a check made, of course, to see that all possible terms of a given count are present.

If a symmetric function has mixed variables, there is no simple prime-counting method which will detect it. If the number of variables is small, one can systematically try, one by one, the transformations described by Mr. Washburn to see whether the original function can be converted to a form in which a symmetric prime count appears. This is at best a laborious process and it becomes prohibitively time-consuming for larger numbers of variables. The map method obviates the need for transformation in the recognition and identification process.

Mr. Karnaugh's contribution speaks for itself. He observed that the identification procedure for functions of more than four variables involved a rather utidy trial and error step, and has described a way to establish the identity of a function by a direct route. It is a most welcome contribution to the subject.

Impulse Ionization and Breakdown in Liquid Dielectrics

G. M. L. SOMMERMAN
FELLOW AIEE

C. J. BUTE
ASSOCIATE MEMBER AIEE

E. L. C. LARSON
ASSOCIATE MEMBER AIEE

Synopsis: A photographic technique was used to measure ionization under impulse voltage in transformer mineral oil and in transformer Askarel (a chlorohydrocarbon). Two electric-field conditions were investigated:

1. The liquid under test surrounded disk electrodes between which were placed impregnated pressboard and dried photographic paper. Above a threshold voltage, Lichtenberg figures occurred in the thin film of liquid under the rounded electrode edge. For mineral-oil immersant, the figures were independent of voltage polarity, while for Askarel immersant, the figures depended markedly on voltage polarity. The chlorine atoms in the Askarel and in carbon tetrachloride are shown to affect the liquid-ionization process and the puncture voltage of Askarel-impregnated pressboard. Chopped-wave tests demonstrate how the discharges develop with time.

2. The pressboard was replaced by a polystyrene sheet and the upper electrode was raised above the sheet to permit the liquid to fill the intervening space. Ionization records were obtained on the photographic paper at threshold stresses which, in most cases, were lower than either the breakdown stress with the polystyrene barrier present or the breakdown stress of the oil without the barrier. The threshold impulse-ionization stress was the same when the photographic paper was under the transparent barrier and not in contact with the stressed oil layer. Threshold-ionization stresses ranged from 400 to 800 kv per centimeter in virtually uniform fields.

THE mechanism of the electrical breakdown of fluids is understood better as a result of analysis of photographic records of presparkover discharges. Thus, for gases, Flegler and Raether[1] photographed discharges between plane-parallel electrodes, and found that a moving space charge was set up. This led to the theory of photoelectric liberation of electrons contributing to the initiating avalanche by ultraviolet radiation. Again, Merrill and von Hippel[2] obtained detailed Lichtenberg figures on photographic plates in gases between rod-plane and similar electrodes. By varying the gas pressure and adding carbon tetrachloride vapor to the test chamber, they studied the various stages of development of discharges and the supression of discharges by electronegative gases.

In like manner, photographs of prebreakdown discharges in liquids are informative. Early workers in this field, such as Marx[3] and Toriyama,[4] made use of at least one pointed or sharp-edged electrode. They showed how the discharges in the liquid develop from the point, and studied the effects of voltage, pressure, and temperature on the extent of the discharge. Toriyama, who used a method of dust figures, further showed that the discharge from a point was twiglike for impulse voltages and bushlike for direct voltages. Wull and Inge,[5] in a study of breakdown of glass photographic plates between disk electrodes immersed in various liquids, found that discharges at the electrode edges affected the photographic emulsion with both impulse and power-frequency voltages.

Our work, which was independent of that of Wull and Inge, began as a result of observations of electrode effects in impulse-breakdown tests of impregnated transformer pressboard. Dakin and Works,[6] in an extensive program of such tests, showed that with transformer-Askarel immersant, breakdown of the pressboard was initiated in the liquid at the edge of the positive disk electrode. While participating in this program in 1948, one of us[7] found that, with mineral-oil immersant, most of the press-board punctures occurred well within the electrode edges. Further analysis[7] suggested that the immersant at the electrode edge was ionizing and that, for mineral oil, this might lead to semiconducting extensions of the electrodes, which would minimize edge failures. Later, the idea evolved that such impulse ionization in the liquid immersant might be observed by placing photographic paper over the sheet insulation, under and beyond the disk electrodes. In 1950, this technique was applied by two of us to a mineral oil and to transformer Askarel.[8] In 1950–51, the third author joined in tests with liquid carbon tetrachloride and chopped-wave tests, and conducted the last phase of the work reported herein on ionization in a liquid layer in series with a solid sheet.[9]

Paper 54-69, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing November 27, 1953.

G. M. L. SOMMERMAN is with the Battelle Memorial Institute, Columbus, Ohio, C. J. BUTE is with North American Aviation, Inc., Los Angeles, Calif., and E. L. C. LARSON is with Otter Tail Power, Fergus Falls, Minn.

The experimental work was performed in the High-Voltage Laboratory of Northwestern University, Evanston, Ill. The assistance of R. P. Bartholomew in constructing the test cells is gratefully acknowledged.

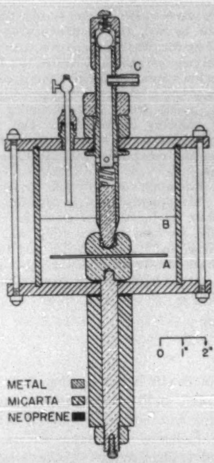

C
B
A
0
1"
2"
METAL ▨
MICARTA ▨
NEOPRENE ■

Fig. 1. Cell for pressboard-breakdown and immersant-ionization tests

A. Pressboard
B. Immersant
C. To vacuum

So far as is known, no other studies have been published of ionization figures for chlorinated liquids, or with procedures and other materials as employed in the current work.

Test Procedures with Pressboard in Immersant

To permit evaluation of the photographic-paper technique, a light-tight insulating container for housing the electrode assembly was constructed, as shown in Fig. 1. The electrodes were disks, 2 inches in diameter, the upper one round edged, and the lower one square edged. This configuration was one of those found by Dakin and Works[6] to yield polarity effects in impulse-breakdown tests on Askarel-impregnated pressboard. The press board used in our tests was 0.059 inch thick, 1.10 specific gravity. Samples of photosensitive paper, 0.007 inch thick, were cut from sheets of grade *F-4* bromide paper. Both the pressboard and the photopaper were vacuum dried and impregnated in separate chambers. The sheets were mounted between the electrodes, with the photopaper emulsion next to the upper electrode. In most cases, the liquid immersant in the cell was not degassed.

Standard $1^1/_2$x40-microsecond impulse-voltage waves were applied to the test cell with a high-voltage surge generator. The polarity of the waves reported herein are referred to the upper, or round-edged, electrode (facing the photopaper).

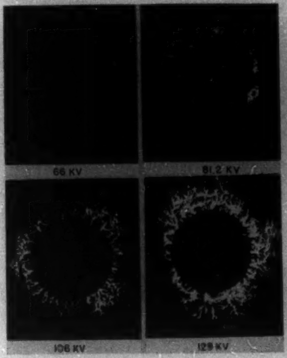

66 KV
81.2 KV
106 KV
129 KV

Fig. 2. Impulse-ionization patterns in transformer-oil immersant at various voltages. Threshold voltage 57.5 kv, breakdown voltage 136 kv

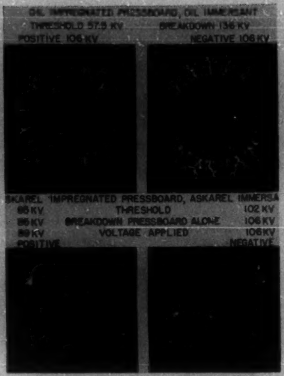

OIL IMPREGNATED PRESSBOARD, OIL IMMERSANT
THRESHOLD 57.5 KV
BREAKDOWN 136 KV
POSITIVE 106 KV
NEGATIVE 106 KV
ASKAREL IMPREGNATED PRESSBOARD, ASKAREL IMMERSANT
66 KV THRESHOLD 102 KV
86 KV BREAKDOWN PRESSBOARD ALONE 106 KV
89 KV VOLTAGE APPLIED 106 KV
POSITIVE
NEGATIVE

Fig. 3. Impulse-ionization patterns in liquid immersants, showing polarity effects

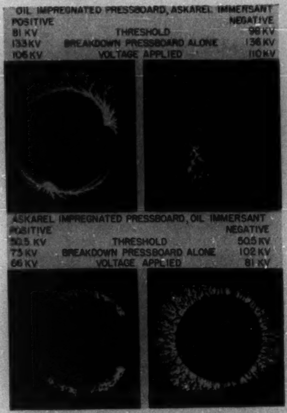

OIL IMPREGNATED PRESSBOARD, ASKAREL IMMERSANT
POSITIVE
NEGATIVE
81 KV THRESHOLD 98 KV
133 KV BREAKDOWN PRESSBOARD ALONE 136 KV
106 KV VOLTAGE APPLIED 110 KV
ASKAREL IMPREGNATED PRESSBOARD, OIL IMMERSANT
POSITIVE
NEGATIVE
50.5 KV THRESHOLD 50.5 KV
75 KV BREAKDOWN PRESSBOARD ALONE 102 KV
66 KV VOLTAGE APPLIED 81 KV

Fig. 4. Impulse-ionization patterns in liquid immersants (pressboard impregnant different from immersant)

For a given combination of impregnant and immersant, the breakdown voltage of the pressboard without photopaper was determined first. Tests were started at a voltage of about 80 per cent of the expected breakdown values; successive impulses increasing by about 3 per cent increments were then applied until puncture occurred.

Following this, samples with photopaper were tested for immersant ionization at various reduced voltages, usually with a single impulse per sample. At values very near the threshold-ionization voltage, several impulses of the same magnitude were sometimes applied to provide a more definite record. This had no effect on the pattern radius. After the photopaper was removed in a dark room, the impregnant was extracted with solvent. After the paper was dry, the photographic record was developed in the standard manner.

The ionization patterns presented in this paper are representative of a large number of photographic records obtained in the investigation. Each threshold and breakdown voltage reported is the average value of several test results for each combination of materials.

Impulse Ionization in Immersant with Pressboard Barrier

Fig. 2 shows photographic records of impulse-ionization patterns in the transformer-mineral-oil immersant surrounding mineral-oil-impregnated pressboard. From the data given, it is seen that the impulse-ionization threshold voltage is about 42 per cent of the pressboard-breakdown voltage. As the voltage is increased, the ionization patterns become defined more clearly and increase in radius. The similarity of these figures to Lichtenberg figures obtained in air is apparent. Tests with degassed mineral-oil immersant gave ionization patterns of the same form and extent as those shown here for the undegassed immersant. Photographs with an external camera show that

Table I. Threshold Impulse-Ionization Voltages of Immersant and Impulse-Breakdown Voltages of 0.059-Inch Pressboard

Materials and Conditions	Ionization and Breakdown Voltages, Kv	
	Positive Polarity	Negative Polarity
Oil pressboard; oil immersant		
Threshold	58	58
Breakdown	136	136
Oil pressboard; Askarel immersant		
Threshold	81	98
Breakdown	133	136
Askarel pressboard; oil immersant		
Threshold	50	50
Breakdown	73	102
Askarel pressboard; Askarel immersant		
Threshold	85	102
Breakdown	85	106

Table II. Approximate Threshold Impulse-Ionization Stresses in Immersant at Round Electrode Edge

Materials	Threshold Stress, Kv/Cm	
	Positive Polarity	Negative Polarity
Oil pressboard; oil immersant	770	770
Oil pressboard; Askarel immersant	540	650
Askarel pressboard; oil immersant	810	810
Askarel pressboard; Askarel immersant	680	820

$$\text{Immersant stress} = \frac{\text{applied voltage}}{\text{sample thickness}} \times \frac{\text{dielectric constant (sample)}}{\text{dielectric constant (immersant)}}$$

Dielectric constants:

Oil	2.25
Askarel	4.5
Oil-impregnated pressboard	4.5
Askarel-impregnated pressboard	5.4

the ionization in the liquid emits light.

The patterns for mineral-oil immersant presented in the upper portion of Fig. 3 show that the characteristics and radius of the pattern are independent of the polarity of the applied voltage. The data indicate that the breakdown voltage of the mineral-oil-impregnated pressboard is also independent of the polarity of the applied voltage. Although the appearance of the patterns suggests that considerable semiconducting shielding effects exist at the electrode edges, breakdown of the pressboard usually occurred at the tangent radius of the upper electrode in these tests.

In these oil-ionization patterns and those presented later, the ionization radius is less in one direction, or two diametrically opposed directions, than it is for the remainder of the figure. These apparent departures in ionization uniformity circumferentially around the electrode are caused by a curling away of the photopaper from the pressboard toward the electrode.

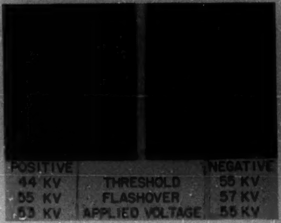


Fig. 5. Impulse-ionization patterns in carbon tetrachloride immersant

Photographic records of ionization patterns in transformer-Askarel immersant and Askarel-impregnated pressboard are shown in the lower portion of Fig. 3. For positive-polarity voltage applied to the round-edged electrode, the pattern indicates circumferential propagation of the ionization in Askarel rather than the nearly radial propagation noted for mineral-oil immersant. For negative-polarity voltages, the figures exhibit restricted bushlike characteristics. In most cases, the ionization originates in only one or two locations as compared with the more nearly uniform ionization with mineral-oil. On the other hand, the maximum density of the ionization is obviously much greater in the Askarel.

Threshold ionization voltages for the Askarel are higher than those for mineral oil. This is to be expected, because the dielectric constant of the Askarel is about twice that of the mineral oil, and so a higher applied voltage is necessary to produce the same electrical stress in the Askarel. With Askarel immersant and impregnant, the breakdown voltages of the pressboard are only slightly higher than the threshold voltages, and are appreciably less than the impulse-breakdown voltages obtained with mineral oil. For the Askarel, both the positive-polarity threshold voltage and the pressboard breakdown voltage are lower than the negative-polarity values.

Data were also obtained on combinations of Askarel immersant and mineral-oil-saturated pressboard, and mineral-oil immersant and Askarel-impregnated press board.[6,7] Typical patterns are shown in Fig. 4. It is seen that the appearance of the patterns depends on the immersant and not on the pressboard impregnant.

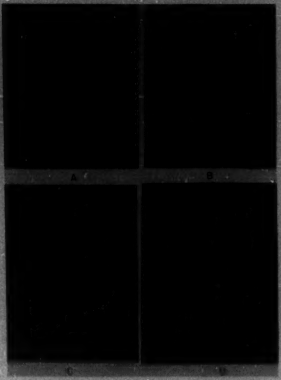


Fig. 6. Ionization patterns in transformer-oil immersant with chopped impulse-voltage waves

Chop times:

- A. 2.5 microseconds
- B. 6 microseconds
- C. 14 microseconds
- D. 30 microseconds

This indicates that diffusion of the impregnant into the immersant was not appreciable.

For these and other combinations of materials, the threshold ionization voltages of the immersant and the breakdown voltages of the pressboard are summarized in Table I. Table II gives approximate values of voltage stress in the immersant where the upper electrode curves away from the pressboard corresponding to threshold ionization. As would be expected, the threshold-voltage stresses depend on the nature of the immersant and not on the impregnant. On the other hand, the pressboard breakdown voltages depend on the impregnant and, in all but one case, not on the immersant. The latter result agrees, for the most part, with results obtained by Wull and Inge[5] on solid materials. In our work, however, the Askarel-impregnated pressboard is a material that does not resist well the effects of ionization in the immersant,[7] especially with positive-polarity voltages. When tested with Askarel immersant, this pressboard breaks down as soon as the threshold ionization voltage of the immersant is reached (85 kv, positive polarity). When tested with mineral-oil immersant, which ionizes at a lower voltage, this pressboard breaks down at only 73 kv, positive polarity.

In order to investigate further the

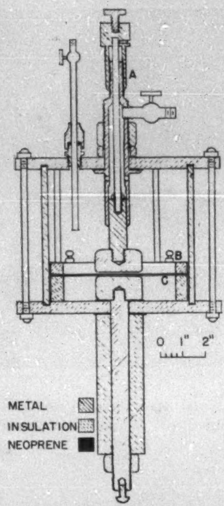


Fig. 7. Cell for liquid-ionization and breakdown tests

- A. Micrometer-spacing adjustment
- B. Barrier and film holder
- C. Insulating barrier and film

effects of structure of chlorinated immersants on impulse-ionization voltage patterns, liquid carbon tetrachloride was tried. In most cases, the pressboard flashed over completely or gave no patterns at all, even when chemically pure carbon tetrachloride was used. A few faint, but well-defined patterns were obtained without flashover, however, and a pair of these is shown in Fig. 5. It is apparent that the positive-polarity figure is very similar to that for the transformer Askarel. The effect of polarity on the threshold ionization voltage is also similar to that for the Askarel. The negative-polarity figure for carbon tetrachloride is not exactly the same as that for the Askarel. This might be expected, since the carbon tetrachloride contains no hydrogen atoms.

Chopped-Wave Tests

In order to evaluate quantitatively the propagation of the ionization streamers in liquid immersants, tests were conducted in which the applied impulse waves were chopped at definite time values. This was done through the use of an auxiliary point gap in parallel with the test cell. Most of the chopped-wave tests were made with $1^1/_2$x300 impulse waves. Four oil-ionization patterns, all obtained at 128 kv and at four different time values, are shown in Fig. 6. At 2.5 microseconds, the pattern is already well developed in the highly stressed region under the curved edge of the upper electrode. At 6 microseconds, the streamers have developed to a greater distance and with more branches. At 14 and 30 microseconds, the branches have grown further, but at a much lower velocity because of the decreased voltage stresses in the medium at the distances penetrated. Worthy of note is the fact that after 2.5 microseconds more new treelets appear to grow out of the electrode edge in between the large trunks and under the main lower branches of the first streamers. These may be similar to the "back-figures" observed in full-wave studies of gas-ionization patterns by Merrill and von Hippel,[2] who attributed such figures to the space-charge current that flows back as the applied-voltage transient subsides.

Analysis of the chopped-wave patterns, together with the full-wave patterns at increasing voltages as shown in Fig. 2, permits computation of streamer velocities in oil at different field intensities. (The difference in streamer lengths at successive chop times divided by the difference in chop times gives the average streamer velocity.) The velocities in oil are but a small fraction of streamer velocities in gases which would be expected from the greater density and greater electron affinity of oil. However, chopped-wave tests on Askarel show that the positive circumferential streamers (similar to those in Figs. 3 and 4) are well developed at about 1.5 microseconds. At threshold, they have not developed at all in slightly less than 1.5 microseconds; yet, with full waves, the positive streamers are much longer than the threshold streamers for oil. Thus, once the positive streamer starts in transformer Askarel, it propagates through adjacent highly stressed regions at a velocity which may reach many centimeters per microsecond. The direction of the streamer is usually perpendicular to the original field. A rapidly moving dense stream or transfer of electrons through the chlorinated liquid is indicated, even though chlorinated materials in the gaseous phase are known to have a high affinity for electrons at low ionization densities.[2]

This rapid electron transfer may also explain the relatively low impulse-breakdown strength of Askarel-impregnated pressboard when ionization occurs in contact with it. Lateral propagation of the positive streamer through highly stressed Askarel in the tortuous capillary paths in the pressboard is possible in a matter of microseconds.

Impulse Ionization in Liquid Layers in a Uniform Field

An additional investigation was conducted to see whether the photographic-paper technique would show the presence of impulse ionization in the liquid layer produced by raising the upper electrode above the surface of the photographic paper on the insulating sheet. In such an arrangement, the electric field in the liquid layer is virtually uniform, computable, and perpendicular to the photopaper.

It was found that Tsikin[10] had investigated a similar geometric arrangement with power-frequency voltages. He reported that discharges were observed at a stress of 250 peak kv per centimeter in an oil film 0.1 millimeter thick and a series glass barrier, with a photographic emulsion in contact with the oil. With the photographic emulsion not in contact with the oil, he reported that the discharges did not occur until the oil stress reached 500 peak kv per centimeter. In order to test for such an effect in our experiments with impulse voltages, it was decided to use transparent polystyrene, 1/32-inch thick, as the insulating barrier instead of the impregnated-paper pressboard used in the earlier experiments.

Another light-tight cell was constructed for these experiments as shown in Fig. 7. It is similar to the first cell except that:

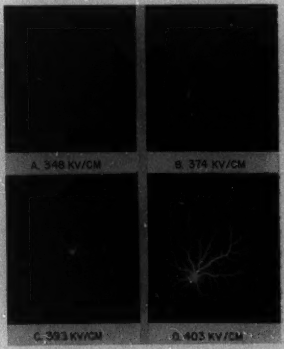

A. 348 KV/CM
B. 374 KV/CM
C. 393 KV/CM
D. 403 KV/CM

Fig. 8. Impulse-ionization patterns in 1/8-inch transformer-oil layer above a polystyrene barrier

Photopaper located over barrier
Negative polarity:
Threshold stress 330 kv/cm
Computed stress at barrier breakdown 410 kv/cm

1. Both electrodes are round edged.
2. The photopaper is clamped to the polystyrene sheet by the plate holder mounted so that the sheet presses against the lower electrode, as shown.
3. The position of the upper electrode above the sheet is fixed by a micrometer adjustment.

Filtered commercial transformer oil was admitted to the cell under vacuum to degas it. The vacuum was broken just prior to the application of voltage. Tests were made with oil-layer thicknesses of 1/8, 1/16, 3/64, and 1/32 inch. Standard $1^1/_2$x40-microsecond voltage waves were applied in all cases.

At relatively low voltages, no effect was observed on either the photographic paper or the voltage oscillograms. At a critical threshold voltage, faint concentrated discharges were observed on the photographic paper. If breakdown of the polystyrene barrier did not occur, no nicks on the voltage oscillograms could be detected in the liquid ionization voltage range. Under most conditions, breakdown of the barrier occurred at voltages greater than the threshold ionization voltages for the liquid; breakdown produced a chopped-wave oscillogram.

For each experimental condition shown in the subsequent figures, the voltage stress in the liquid layer was computed from the following relationship for a 3-layer dielectric.

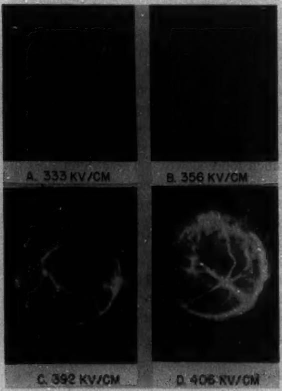


Fig. 9. Impulse-ionization patterns in 1/8-inch transformer-oil layer above a polystyrene barrier

Photopaper located under barrier
Negative polarity:
Threshold stress 330 kv/cm
Computed stress at barrier breakdown 410 kv/cm

$$E_1 = V_T/[d_1 + (\epsilon_1 d_2/\epsilon_2) + (\epsilon_1 d_3/\epsilon_3)]$$

where

V_T = the total applied voltage
d_1, d_2, d_3 = the thicknesses of the three layers
ϵ_1, ϵ_2, ϵ_3 = the dielectric constants of the three layers

Fig. 8 shows ionization patterns at four applied voltages for the 1/8-inch-thick oil layer with the photopaper in contact with the oil. Each pattern consists of only one ionization nucleus, with streamers radiating in a characteristic dendritic pattern, the radius of which increases with voltage. Fig. 9 shows results for similar tests except that here the photopaper is under the polystyrene barrier (emulsion not in contact with the oil layer, but still facing the oil layer). The results are similar to those illustrated in Fig. 8, except that the patterns are more diffuse in character because of the intervening sheet.

The patterns for the 1/16-inch oil layer are similar to those for the 1/8-inch oil layer. In both cases, the patterns were obtainable only with negative-polarity applied voltages before breakdown of the polystyrene barrier occurred. For the 3/64-inch oil layer, more than one ionization nucleus is usually observed, at both polarities, and the pattern radii are small, as shown in Fig. 10. For the 1/32-inch oil layer, numerous pointed or small-radius starlike patterns are observed, as shown in Fig. 11. The occasional corona-like figure (about 0.1-inch diameter, actual size), is interesting.

Both Figs. 9 and 10 include patterns recorded on photopaper mounted under the polystyrene barrier. For all four oil-layer thicknesses investigated, the threshold ionization values for emulsion in contact with the oil agreed within 3 per cent with those for emulsion not in contact with the oil. It should be noted that these results are obtained under experimental conditions similar to those of the pressboard studies; namely, well-dried photopaper and impulse voltages.

In many of the patterns obtained with the photopaper in contact with the oil, Figs. 8, 10, 11, thin circular lines are evident. These are scratches caused by contact of the top electrode with the surface of the photopaper when the micrometer zero setting for the liquid-layer spacing was made; they are useful in indicating the location of the electrode edge.

In the patterns obtained with the photopaper under the polystyrene barrier, Figs. 9, 10, diffuse circular halos are evident when the voltage is substantially higher than the ionization threshold. The halos are not observed at, or slightly above, threshold levels. An explanation of this is the following: as the voltage is increased above threshold, the ionization in the liquid layer extends farther from the electrode center, and may extend beyond the electrode edge, see Fig. 8. Thus, the higher the voltage, the more charge is sprayed on the upper surface of the barrier, and the more closely the potential of the upper barrier surface approaches the upper electrode potential. This increase in potential of the upper barrier surface increases the stress in the barrier and in the liquid layer at the curved edge of the lower electrode. The liquid layer at the lower electrode edge thus ionizes sufficiently to affect the photographic emulsion through the liquid-soaked photopaper, giving a diffuse halo. With lower voltages, ionization in the main liquid layer is limited, little charge is sprayed on the barrier, and the stress in the liquid under the barrier does not reach the ionization level.

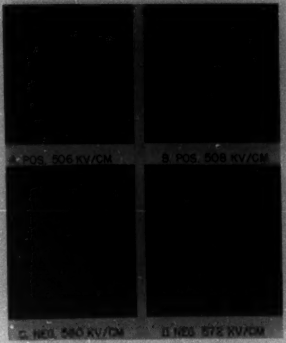


Fig. 10. Impulse-ionization patterns in 3/64-inch transformer-oil layer above a polystyrene barrier

A and C, Photopaper under barrier
B and D, Photopaper over barrier
Threshold stresses: 493 kv/cm, positive; 453 kv/cm, negative
Computed stresses at barrier breakdown: 532 kv/cm, positive; 586 kv/cm, negative

Fig. 12 shows the threshold ionization voltage stresses in the oil plotted against oil-layer thickness for both voltage polarities. The threshold values increase as the thickness decreases. The extrapolated values at thicknesses of a few mils are consistent with the 800 kv per centimeter (kv/cm) values computed from the results of the pressboard-in-immersant studies, Table II.

The relationship of barrier-breakdown

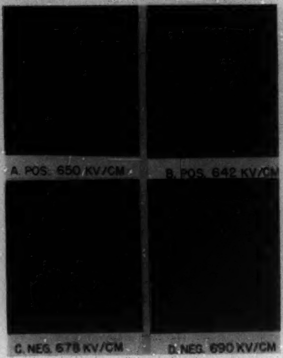
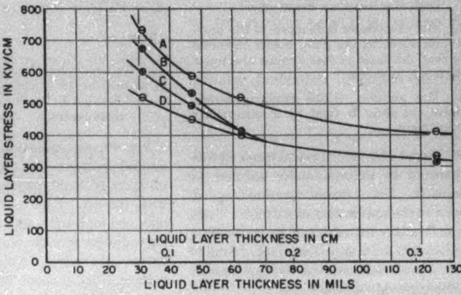


Fig. 11. Impulse-ionization patterns in 1/32-inch transformer-oil layer above a polystyrene barrier

A and C, Photopaper under barrier
B and D, Photopaper over barrier
Threshold stresses: 600 kv/cm, positive; 520 kv/cm, negative
Computed stresses at barrier breakdown: 670 kv/cm, positive; 730 kv/cm, negative

Fig. 12 (right). Impulse - ionization stress versus oil-layer thickness. Computed oil-layer stress corresponding to breakdown of polystyrene barrier

A. Negative-polarity breakdown
B. Positive-polarity breakdown
C. Positive-polarity threshold ionization
D. Negative-polarity threshold ionization



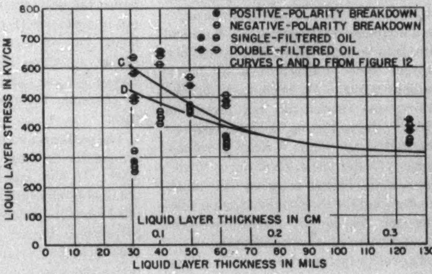


Fig. 13 (right). Impulse - breakdown stress versus oil-layer thickness, without barrier

voltage to threshold voltage at various oil-layer thicknesses is shown by the curves for computed oil-layer stresses corresponding to barrier breakdown in Fig. 12. The actual oil-layer stress at barrier breakdown is probably less than indicated by negative-polarity curve *A*, because of the formation of charged areas on the upper barrier surface discussed in the foregoing.

Impulse-breakdown tests were made on the degassed oil alone (without the barrier), using the same electrodes. The results are given in Fig. 13. Even with single-filtered oil the breakdown stresses are low and erratic, particularly for oil-layer thicknesses of 1/16 inch and less. Using double-filtered oil, breakdown stresses are higher and more uniform; they increase with decreasing oil-layer thicknesses down to 0.040 inch. For the 0.030-inch oil layer, results were low and erratic even with double- and triple-filtered oil. In general, these oil-breakdown stresses are higher than the impulse-ionization threshold stresses, which are redrawn in Fig. 13 for comparison. This is partly explained by the fact that the liquid does not break down until after about four microseconds on the impulse-voltage wave. Therefore, if the voltage is reduced slightly below the breakdown point, ionization in the liquid may still occur at the crest-voltage value.

Similar experiments were conducted with transformer Askarel and a glass barrier. In nearly all cases, breakdown of the barrier occurred before ionization was photographed. A few faint photographic records were obtained at voltages very slightly less than those causing barrier breakdown.

Summary and Conclusions

Impulse ionizaton in liquids may be detected by photographic paper mounted in contact with the ionization. The effect is caused by light emitted by the ionization in the liquid. The presence of dry photographic paper does not appreciably change the impulse ionization being measured. In degassed transformer oil, impulse ionization is usually detected at voltage stresses less than those causing complete breakdown of the oil in essentially uniform fields.

The threshold impulse-ionization stress increases as the thickness of the liquid layer decreases. In these tests it approached 800 kv/cm for both transformer mineral oil and transformer Askarel. In a diverging field, there is no appreciable polarity effect on the threshold impulse-ionization stress for mineral oil, but for transformer Askarel and carbon tetrachloride, the ionization occurs at a lower stress when the field diverges from a positive electrode toward a dielectric barrier. The mechanism of propagation of the impulse ionization depends on the chemical composition of the liquid. For a hydrocarbon oil numerous streamers propagate relatively slowly in treelike patterns in the general direction of the original field for both polarities. For a chlorohydrocarbon liquid, positive streamers propagate more rapidly from a very few ionization nuclei and laterally to the field; negative streamers propagate much more slowly. These differences in streamer propagation are reflected in corresponding differences in impulse-breakdown strengths of pressboard impregnated with the liquids, if the pressboard is exposed to ionization. The photographic method for detecting ionization may be employed to determine the location of weak spots in insulated structures.

References

1. **Investigation of Gas-Discharge Processes with the Cloud Chamber, E. Flegler, H. Raether.** *Zeitschrift für technische Physik*, Leipzig, Germany, vol. 16, 1935, pp. 435–40.

2. **The Atomphysical Interpretation of Lichtenberg Figures and Their Application to the Study of Gas Discharge Phenomena, F. H. Merrill, A. von Hippel.** *Journal of Applied Physics*, New York, N. Y., vol. 10, 1939, pp. 873–87.

3. **The Electric Breakdown of Air in Nonhomogeneous Fields with Various Types of**

Voltages, **B. Marx.** *Elektrotechnische Zeitschrift*, Wuppertal-Elberfeld, Germany, vol. 51, 1930, pp. 1161–65 (photograph of discharge in oil attributed to H. Staack).

4. Lichtenberg Figures in Gases and Liquids, **Y. Toriyama.** *Proceedings of the Physico-Mathematical Society of Japan*, Tokyo, Japan, vol. 14 1932, pp. 430–41

5. The Nature of the Edge Effect, **B. Wull, L. Inge.** *Zhurnal Tekhnicheskoi Fisiki*, Moscow, U.S.S.R., vol. 1, 1931, pp. 129–35.

6. Impulse Dielectric Strength Characteristics of Liquid-Impregnated Pressboard, **T. W. Dakin, C. N. Works.** *AIEE Transactions*, vol. 71, pt. I, Nov. 1952, pp. 321–25.

7. Discussion by **G. M. L. Sommerman.** *Ibid.* p 325.

8. A Study of Edge Phenomena in Electrical Breakdown Tests on Solid Insulation, **E. L. C. Larson.** *Thesis*, Northwestern University, Evanston, Ill., 1950.

9. Impulse Ionization and Breakdown of Liquids in Uniform Fields, **C. J. Bute.** *Thesis*, Northwestern University, Evanston, Ill., 1951.

10. Determination of the Intensity of the Electric Field Which Initiates Discharge in Thin Layers of Liquid Dielectrics, **A. Tsikin.** *Technical Physics of the U.S.S.R.*, Moscow, U.S.S.R., vol. 3-II, 1936, pp. 947–55.

◆

Discussion

T. W. Dakin (Westinghouse Research Laboratories, East Pittsburgh, Pa.): The authors have reported the results of a very interesting and significant investigation of a phenomenon associated with the dielectric breakdown of sheet material. It is probably not generally appreciated that dielectric breakdown of solid insulation in oil is almost always preceded by a discharge in the oil at the conductor edge. This confirmation of this discharge phenomenon, even with impulse tests of short duration, is certainly going to be helpful in advancing our understanding of dielectric breakdown in composite systems of this sort.

I believe that the results of this investigation also have a bearing on the long-time dielectric strength of pressboard in oil. In a paper by Dakin, Philofsky, and Divens[1] it is pointed out that the long-time application of alternating voltages under conditions of surface or internal discharges reduces the effective dielectric strength until failure occurs. The lower voltage limit at which this effect of discharge attack occurs seems to be the corona or discharge inception voltage. The authors of the present paper have shown in the case of oil-impregnated pressboard in Table I that the threshold voltage of the discharges with an 0.059-inch pressboard is 58 kv as compared to an impulse breakdown voltage of 136 kv crest. Thus we might expect that with repeated discharges, such as would occur with alternating voltage, attack on the pressboard would occur at crest voltages above 59 kv. In recent tests at our Laboratory on pressboard of comparable thickness we obtained the results shown in Table III.

The data in Table III indicate that the discharges which the authors detected occurred at a voltage corresponding to breakdown on a 2-kv per minute step test with 60-cycle voltages. Longer time 60-cycle tests, however, showed lower breakdown voltages. It is therefore probable that discharges occurred at lower voltages but were not detected by the photographic method used by the authors. Perhaps the authors would like to comment on the sensitivity of the photographic method of detecting these discharges.

Table III

Method of Voltage Application	KV Crest Breakdown
$1^{1}/_{2}$x40-microsecond impulse	141
60 cycle—2 kv per second rise	88
60 cycle—2 kv steps per minute	62.4
60 cycle—2 kv steps per 10 minutes	49.5

The authors' result, which indicates that in the case of Askarel-impregnated pressboard the discharge threshold voltage was very near to the breakdown voltage, correlates with the low impulse ratio for this material reported by Dakin and Works (see ref. 6 of the paper). With this material the long-time 60-cycle crest breakdown voltage is only a little lower than the impulse level because the discharges in the liquid do not occur at a lower voltage.

Reference

1. Effect of Electric Discharges on the Breakdown of Solid Insulation, **T. W. Dakin, H. M. Philofsky, W. C. Divens.** *AIEE Transactions*, vol. 73, pt. I, May 1954, pp. 155–62.

C. N. Works (Westinghouse Research Laboratories, East Pittsburgh, Pa.): The authors have described a practical method of obtaining patterns on photographic paper of the corona which results when impregnated pressboard in a liquid immersant is subjected to impulse voltage of sufficient magnitude. Impregnated pressboard in a liquid immersant is used as a dielectric in practically all large power transformers built. The fine photographs of ionization or corona patterns obtained by these authors are a worthy contribution toward a better understanding of the mechanism of corona discharge and dielectric breakdown that result from the application of impulse voltages.

I had the pleasure of being associated with Dr. Sommerman in a program of impulse breakdown tests of impregnated pressboard (see ref. 7 of the paper). Soon after Dr. Sommerman returned to Northwestern University, some unreported tests were made on oil-impregnated and Askarel-impregnated pressboard immersed in both Askarel and oil as well as in oil containing acid and other contaminants. Results similar to those given in Table I were obtained and, in general, I concur with the conclusions reached by the authors in the section entitled "Impulse Ionization in Immersant with Pressboard Barrier." However, attention should be called to the following facts. When 1/16-inch Askarel-impregnated pressboard immersed in Askarel was tested with a $1^{1}/_{2}$x40-microsecond impulse wave, breakdown occurred at 85 kv positive. For 1/16-inch Askarel-impregnated pressboard immersed in oil, breakdown occurred at only 73 kv positive; see Table I. This is a considerable reduction in dielectric strength. This was explained by the fact that the threshold corona voltage for the case of Askarel-impregnated pressboard immersed in oil is considerably lower than for the same immersed in Askarel, and that this pressboard does not resist well the effects of corona in the immersant. I accept the fact that the threshold corona voltage of Askarel-impregnated pressboard immersed in oil is lower than that for the same immersed in Askarel. However, perhaps the statement that Askarel-impregnated pressboard does not resist well the effects of corona in the immersant should be reexamined.

When Askarel-impregnated pressboard is immersed in oil, the stress in the oil is greater than in any other case; see Table II. The impulse dielectric breakdown voltage of Askarel-impregnated pressboard is lower than for oil-impregnated pressboard, Table I and reference 6 of paper. As an impulse voltage is applied to Askarel-impregnated pressboard immersed in oil, severe corona discharge takes place in the oil. In the absence of an exact knowledge of the breakdown process, it may be suggested that the effect of this discharge is to extend the electrodes a short distance into the pressboard and reduce its effective thickness. Impregnated pressboard is not a homogeneous solid, but consists of fibers surrounded by the liquid impregnant. The separation between fibers is great in terms of atomic dimensions. If the effect of corona in the oil immersant is to reduce the effective thickness of Askarel-impregnated pressboard, then as the sheet thickness is increased this effect should be proportionately less.

I have made breakdown tests on 1/8-inch pressboard, double the thickness used in the paper. Askarel-impregnated pressboard immersed in Askarel had an average positive impulse breakdown voltage of 132 kv. Askarel-impregnated pressboard immersed in oil had an average breakdown voltage of about 130 kv, only a slight reduction as expected. Therefore Askarel-impregnated pressboard immersed in oil is not necessarily less corona resistant than other pressboard and this case is not an exception to the general rule as stated by the authors, but is rather in agreement with their statement made in the paper that impulse breakdown voltages for pressboard depend on the impregnant and not on the immersant.

To avoid confusion with the term ionization potential as used by the physicist in connection with spectroscopy, it is suggested that the term threshold impulse discharge voltage be used in place of threshold impulse-ionization voltage.

J. G. Anderson and **T. W. Liao** (General Electric Company, Pittsfield, Mass.): The use of photographic methods to study streamer formation, creepage, and puncture of materials in liquid dielectrics is a very promising field of investigation. By such a timely paper the authors have made an interesting contribution to this subject.

At the same time that the authors were making the studies reported in this paper,

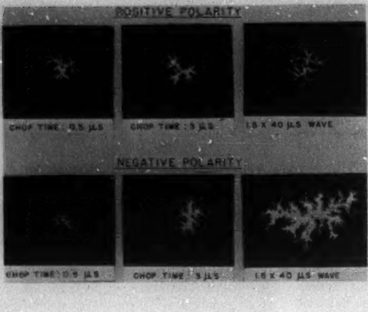


Fig. 14 (left). Effect of duration of impulse wave on Lichtenberg figures

All impulses at 100 kv maximum at time of chop. Wave chopped through damping resistance to eliminate polarity reversal on chop. All film emulsion side up in oil on 1/16-inch kraft board barrier. 1/16-inch-diameter brass electrode on film surface

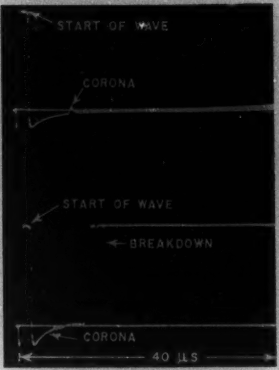


Fig. 15. Oscillograms showing corona prior to breakdown from a pair of 5-inch-diameter Rogowski electrodes spaced 1/2 inch apart and covered by 2.5 mils polyethylene film

Upper—taken at 440 kv, negative 1.5x40-microsecond wave

Lower—taken at 480 kv, negative 1.5x40-microsecond wave

we were using somewhat similar methods in an investigation of streamer processes on insulating surfaces in oil. As a result, we have also observed some of the phenomena reported here.

Our work also shows that the lack of an edge effect during impulse breakdown of insulating board in oil is due to streamers from the electrode edges moving out over the surface of the barrier and grading off the strong fields at the edge. Here is a case where impulse corona on an insulating surface is actually helpful instead of harmful, in that its reduction of strong fields at electrode edges may greatly reduce the possibility of puncture at the edge. However, for weaker insulating materials such as ordinary glass, the glass may puncture at the electrode edge. Hence the lack of an edge effect will occur principally for those materials whose impulse dielectric strength is higher than the strength of the immersant.

Some observations of streamer propagation by chopped waves were also made by us. Fig. 14 shows impulse Lichtenberg figures of a rod-to-plane gap in oil. Here a 1/16-inch kraft board sheet rested directly on a ground plane. On the top surface of the kraft board a sheet of photographic film was laid emulsion side up and a 1/16-inch brass rod line electrode pressed directly against the emulsion. The effect of chopping time on propagation is more pronounced for negative polarity than for positive polarity. It would be interesting to know the polarity of the figures in Fig. 6.

In addition, certain fine points of the test procedure are of interest to us. Did the authors find that the opposite polarity oscillation normally occurring when a wave is chopped had any observable influence on their chopped-wave figure patterns? Also, what is their estimate of the effect that streamers on the opposite side of the pressboard may have had on the pattern of the streamers on the photographic paper. Presumably streamers were occurring simultaneously on both sides of the pressboard barrier.

It is stated in the synopsis, conclusion, and the discussion referring to Figs. 12 and 13 that the threshold ionization stresses were in most cases lower than the breakdown stresses with or without a barrier. It is readily visualized that, in the case with a barrier, ionization can exist in oil in earlier stages without breaking down the barrier.

Results obtained at the General Electric High Voltage Laboratory in 1950 indicated that such a case can exist even though the barrier is as thin as 2.5 mils of polyethylene film covering a pair of 5-inch-diameter Rogowski electrodes (uniform field) spaced 1/2 inch apart.

Average breakdown stress of covered electrodes..............920 kv per inch
Average threshold ionization stress of covered electrodes...806 kv per inch
Average breakdown stress of bare electrodes....................814 kv per inch

The polyethylene covers were vacuum-treated and pulled tight over the electrode under oil by means of a drawstring sewed in the covers.

The ionization or corona was detected by an impulse corona detector described in reference 1 and the typical oscillograms are shown in Fig. 15. However, no corona prior to breakdown level was detected with the bare electrodes. In fact, corona and breakdown levels are the same even for rod gaps (nonuniform field) below 1-inch spacing.

Reference

1. Impulse Corona Detection, Measurement of Intensity, and Damage Produced, J. H. Hagenguth, T. W. Liao. *AIEE Transactions*, vol. 71, pt. III, 1952, pp. 461–65.

G. M. L. Sommerman, C. J. Bute, and **E. L. C. Larson:** It is a pleasure to receive these stimulating discussions of our paper, and to note the agreement of corresponding results obtained in different laboratories.

We concur with Dr. Dakin's observations that the results of our investigation have a bearing on the long-time dielectric strength of liquid-immersed pressboard. The impulse patterns indicate that the curve of ionization current density versus electric stress for Askarel rises very sharply from virtually zero at a critical threshold stress. The nearly unity impulse ratio noted for Askarel-impregnated pressboard indicates that there is little effect of time of stress application on this ionization curve. For mineral oil, the ionization curve apparently rises more gradually, at least for thin oil films. Evidently, there is more effect of time of stress application on the ionization curve for oil, as shown by the somewhat lower value of long-time breakdown voltage compared with the impulse-ionization threshold voltage for oil. Other investigations of 60-cycle ionization threshold stress indicate values less than 600 peak kv/cm for thin films of oil.

The preceding discussion shows that the voltage sensitivity of a method of detecting ionization depends on the type of liquid and the thickness of the liquid film. For Askarel, the photographic method is highly sensitive. For thick films of mineral oil, the repeatability of threshold ionization levels is within about 5 per cent. For thin oil films, the repeatability of results is poorest, nearly 10 per cent.

Mr. Works' discussion on the breakdown process in impregnated pressboard is appreciated. We agree that increasing the thickness of the pressboard reduces the effect of the immersant ionization on the impulse breakdown voltage. In the case of oil immersant and impregnant, increasing the pressboard thickness from 1/16 inch to 1/8 inch gives an increase in resistance to ionization at the electrode edges which is sufficient nearly to eliminate edge failures of the pressboard. With 1/8-inch Askarel-impregnated pressboard, however, edge failures still occur with either Askarel or oil immersant unless the pressboard surface is protected. One method of protection is to use shielding, as shown by Dakin and Works (ref. 6 of the paper). There is a little evidence in our data that placing a thin sheet of oil-impregnated paper over the Askarel-impregnated pressboard under the positive electrode gives an increase in breakdown voltage which is larger than when an Askarel-impregnated sheet of paper is used. For these and other reasons we hold the view that Askarel-impregnated fibrous material does not resist well the effects of impulse ionization at the positive electrode. Apparently the ionization streamer in Askarel is not readily halted by the first fibers encountered. By a propagation mechanism probably involving a photoelectric process, the stream may move through the Askarel laterally to the original field and around the fibers, then parallel to the field through the liquid between fibers until the next barrier fiber is reached, whereupon the process is repeated. Such a mechanism can account for the short times observed for complete disruption of such pressboard at moderate impulse stresses.

The results given by Mr. Anderson and Dr. Liao in Fig. 14 are similar to those that we obtained with a 1/8-inch oil layer above a barrier. Here, only one ionization streamer developed (see Fig. 8). For negative polarity, electrons moved readily over the barrier surface. For positive polarity, the thin polystyrene barrier punctured almost immediately, at the threshold voltage, indicating that charges did not move far over the barrier surface, thus causing a stress concentration in the barrier.

In reply to the questions, the polarity of the patterns in Fig. 6 was negative. Similar positive-polarity patterns were obtained with this electrode system. A damping resistor was not used directly in series with the chopping gap, but an auxiliary resistance-capacitance wave-shaping circuit was used, and the resulting oscillation was highly damped. The variation of pattern radius with chopped waves was consistent with the pattern radius for full waves. Photopaper applied to the under side of the barrier showed that streamers did exist there, as expected. These streamers raised the stress on the upper side (at the tangent radius of the curved electrode edge) only slightly when ionization was just starting on the upper side. At higher voltages, the two sets of streamers lengthen because of the radial component of the electric field.

The last results given by Anderson and Liao for two barriers are very interesting. The differences in results are quite large considering the extreme thinness of the barriers used.

It would be interesting to compare the threshold levels obtained with the impulse corona-current detector with those obtained with the photographic method, using oil and the same electrodes in each case. From the results already presented, agreement appears to be good.

Effect of Electric Discharges on the Breakdown of Solid Insulation

T. W. DAKIN
ASSOCIATE MEMBER AIEE

H. M. PHILOFSKY
ASSOCIATE MEMBER AIEE

W. C. DIVENS
STUDENT MEMBER AIEE

THE presence of corona at the surface of electric insulation has long been recognized[1] as a possibly damaging situation. An attempt is usually made by electric apparatus designers to avoid corona at operating voltages or to minimize it. This is done by impregnation with liquids, waxes, or resins or by grading the stress away from edges where corona would occur. Sometimes it is done with the primary purpose of eliminating radio interference rather than protecting the insulation.

Corona is a very appreciable factor in affecting the dielectric strength of solid insulation as it is measured in practical tests. This is true because intense corona occurs at electrode edges previous to the breakdown of the solid, unless considerable care is taken to avoid it. The corona is very apparent with a-c breakdown tests in air. Corona also occurs and may be even more serious in the case of breakdown tests in mineral oil, although it may not be so apparent.

Corona can be described as a partial breakdown of insulation due to concentration of electric stress in divergent electric fields at sharp edges, fine wires, etc., or due to increased stress or lower dielectric strength (or both) in one component of different insulating materials in series. A common situation is a series air gap which is subject to a higher electric stress in proportion to the dielectric constant of the solid in series with it. The dielectric strength of the air gap is also lower than that of the solid. Both factors lead to initial breakdown of the air gap at voltages considerably below those required to puncture the solid. Such air gaps occur in porous or incompletely impregnated insulation, at the surface of insulation under conductors, or at edges of electrodes or conductors in contact with insulation. Similar gaps in oil in series with solid dielectric may also break down initially before the solid fails. Corona or discharge starting voltages in oil are usually much higher than in air, of course, due to the higher dielectric constant and strength of the oil. Breakdown of such series gaps in air, or oil, may produce higher stresses on the solid and also erode the surface of the solid reducing its thickness, in time, to such a value that puncture occurs.

Corona is probably the most important factor in decreasing the dielectric strength of solids with time of alternating voltage application, but it is not the intention of this paper to imply that it is the only factor. Dielectric heating due to a-c conductivity is with some materials the major factor in determining the long time a-c dielectric strength of solid dielectrics.

It is the purpose of this paper to outline a number of the factors affecting the breakdown of dielectrics with corona and to describe the variation of a-c dielectric strength with time of voltage application for a variety of dielectrics.

Minimum Corona Voltages in Air

The threshold voltage for corona is, of course, the minimum voltage at which it can affect the solid insulation. Corona at the surface of isolated conductors away from dielectric surfaces is only of secondary effect on the insulation. It may produce ozone or oxides of nitrogen which diffuse to near-by insulation attacking it chemically. This paper will not be concerned with such secondary attack, but will be concerned only with corona occurring directly between the dielectric and a conductor or between two dielectric surfaces with, in every case, an appreciable component of the electric field normal to the dielectric surface.

Corona threshold voltages in air are determined by the electric field, the pressure of the air, and the thickness of the gap (Paschen's law). In Fig. 1 are shown the corona threshold voltages for an electrode system consisting of two 1/2-inch-diameter square edge electrodes on opposite sides of a plane sheet of insulation. Since the electric field in the air at the electrode edges should be proportional to the dielectric constant of the solid dielectric, the corona threshold voltage should be a function of the ratio of the dielectric thickness to the dielectric constant. That this is approximately so is demonstrated in Fig. 2, where the data obtained on a variety of insulating materials as shown in Fig. 1 over a range of thicknesses from 1 to 250 mils fall almost on a single line when plotted as a function of t/ϵ'. The dielectric constants of these materials range from 2.2 for polyethylene to 8.3 for

---

Paper 54-70, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted May 13, 1953; made available for printing November 27, 1953.

T. W. DAKIN, H. M. PHILOFSKY, and W. C. DIVENS are with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

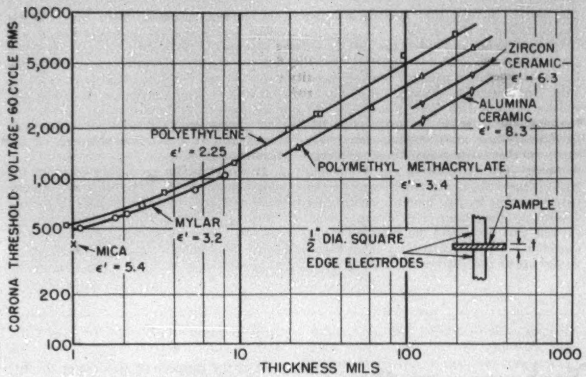


Fig. 1. Corona threshold voltages on dielectric surfaces

the alumina ceramic. The effect of various electrode shapes on the corona starting voltage is demonstrated in Fig. 3 for a single material—polyethylene. By reference to the principle shown in Fig. 2, it is possible to estimate the threshold voltage for other materials with similar electrodes to those of Fig. 3.

The data of Figs. 1, 2, and 3 represent corona offset voltages at 74 centimeters of mercury air pressure. The offset voltage is the voltage at which corona, once started, ceases as the voltage is slowly lowered. This is considered a more consistent and reliable measurement than the onset voltage. The offset voltage is usually a little lower than the onset. The difference is greater with lower pressures and larger electrode spacings.

In Fig. 3, it can be noted that the corona threshold voltage decreases as the divergence of the electric field is increased by increasing the dissymmetry of the electrodes. The variations of threshold voltage with the different electrodes of Fig. 3 are about what would be expected from a consideration of the relative electric field intensity at the electrode edge.

Relation of Breakdown Time to Voltage Gradient

The dependence of the dielectric strength on the time of voltage application has been studied by a number of investigators,[2-7] some of whom have developed empirical equations for this dependence. The relation of the voltage endurance of insulation to corona has not, however, always been appreciated. Recent papers by Brodhun and Perkins[12,13] have reemphasized the importance of corona and pointed out that all organic materials are susceptible to corona erosion.

An example of the relation of dielectric strength to breakdown time is given in Fig. 4 for a 5-mil nylon film. The dielectric strength decreases steadily with increasing time of voltage application, but appears to approach asymptotically the gradient at which corona starts, for very long times of voltage application.

The data plotted in Fig. 4 and in most of the other curves in this paper were, with the exceptions noted, determined with a very simple electrode system; a 1/2-inch cylindrical brass electrode, with flat end and square edge opposite a large plane electrode. The electrodes were initially polished and placed tightly against the flat dielectric sheet under test. The voltage was raised quickly, except in the very short-time tests, to the test voltage and maintained constant until failure occurred. In the long-time tests, usually a number of samples were tested in parallel with a single fuse or circuit breaker. When a sample failed, the voltage remained off until an operator noted the failure, recorded the time, disconnected the failed sample, and reapplied the voltage to the remaining samples. A running time meter in parallel with the primary of the high-voltage transformers recorded the elapsed time of voltage application. This rather simple technique kept the control circuitry to a minimum.

This test procedure, it will be noted, applied the voltage discontinuously. It was determined in a number of tests that discontinuous application of voltage in the longer time tests resulted in the same breakdown times as continuous voltage application. It should also be noted that conventional dielectric strength step tests, where the voltage is increased in intervals, being held for specified times at each interval until failure occurs, gives about the same information as the procedure used here. The step test, however, is believed to be less accurate, since some corona deterioration occurs at lower steps and the time of failure at the terminal step is often not recorded.

The voltage gradient, which is the ordinate of Fig. 4 and later figures and is often referred to throughout the paper, is the average gradient, the ratio of the applied voltage to the electrode separation (usually the thickness of the dielectric under test).

The decreasing dielectric strength with time of alternating voltage application begins at the first cycle, as is shown in

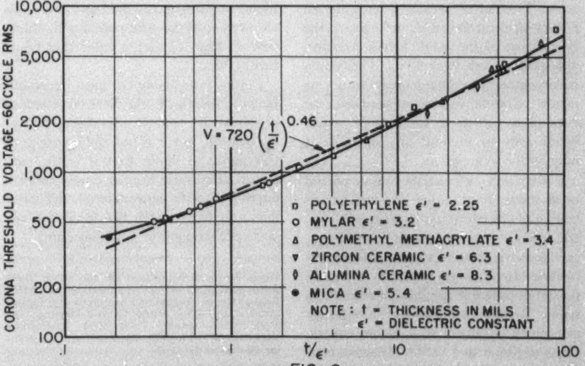


FIG. 2

Fig. 2. Corona threshold voltages on dielectric surfaces

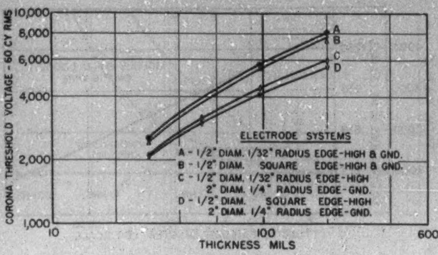


Fig. 3. Corona threshold voltages on polyethylene

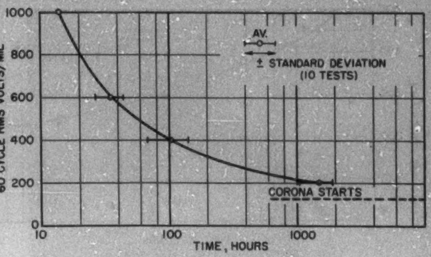


Fig. 4. Voltage endurance, 5-mil nylon

Fig. 5 for the case of 2-mil mylar (polyethylene terephthalate) film. For this material the dielectric strength has been measured over a range of times of voltage application of almost 10^8. All of these tests were made in air. To determine the failure times of about a minute or less, the voltage was applied to the samples with a high-voltage switch from an energized transformer and the time of voltage application measured with a recorder capable of recording a 60-cycle wave. The high-voltage switch technique was necessary to avoid the overvoltage surge often noted when switching was done in the primary.

Difficulty was occasionally encountered in the higher voltage tests because of surface flash around. This was relieved by replacing the larger plane electrode with one of smaller size. Most of the higher voltage tests (darkened circles in Fig. 5) were made with a 1/2-inch square edge electrode opposite a 1-inch round edge electrode. The lower voltage tests (open circles) were made with the ground electrode 2 inches or greater in diameter. At 3,800 volts per mil tests with both electrodes were made and appear to give almost the same results. Some tests were also made with two 1/2-inch-diameter square edge electrodes (half-darkened circles).

Immersing the solid insulation in oil appears to decrease the time to failure very greatly unless the sample is vacuum-treated. This is illustrated by the solid triangles of Fig. 5 for 8-mil mylar. These tests were made by simply immersing the mylar film in oil. The oil appeared to wet the film well, but there must have been microscopic air bubbles or an invisible air film left. When the immersed insulation was thoroughly vacuum-treated, the times to failure were longer However, such tests were sometimes erratically short, possibly as a result of air bubbles retained on the electrodes from previous tests. (The electrodes were not vacuum-treated between tests on the film.) The short times to failure in oil must be due to concentration of the corona at microscopic air bubbles. Even in the absence of bubbles initially it is probable that, if the dielectric strength of the oil is exceeded locally, the local breakdown of the oil will produce gas, forming bubbles.

In Fig. 6 are illustrated data obtained on a thick glass melamine laminate tested in oil. It should be noted that the times to failure here are very short for rather low gradients compared to the times to failure of the mylar and nylon films tested in air. Very likely the concentration of corona at microscopic surface bubbles accelerated the failure as it did in the case of the mylar immersed in oil. It is possible that corona in microscopic internal voids in this laminate also speeded the breakdown. Skow and Belsterling[3] have published data on the voltage endurance of a number of other laminates tested in oil. They did not discuss the reason for the decreasing dielectric strength with time, but it was probably due to corona.

The possibility of breakdown due to thermal rise (dielectric heating) in the case of the melamine laminate was considered, since such failure might show a similar time effect. It was discarded when it was found that intermittent application of voltage, with long periods of rest between, produced failure in almost the same cumulated time as continuous application of voltage. Furthermore, an examination of the surface of samples which had been exposed to voltage for several minutes, but had not yet failed, showed erosion pits, as illustrated in Fig. 7. This microscope photograph of a laminate surface shows the rapid and permanent damage produced by corona (under oil in this case).

It is not necessary that the corona be from a metal electrode, as is shown in Fig. 8 where the corona occurred between the surfaces of the enamel on twisted pairs of no. 16 enameled wire. The time to breakdown in this case is quite com-

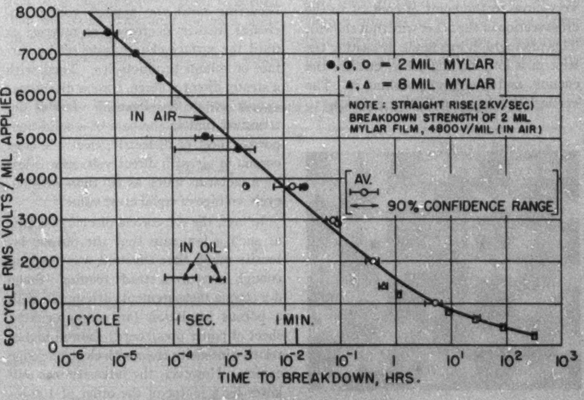


Fig. 5. Voltage endurance, mylar film

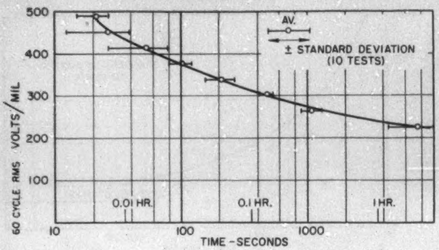


Fig. 6. Voltage endurance, 133-mil glass melamine laminate in oil

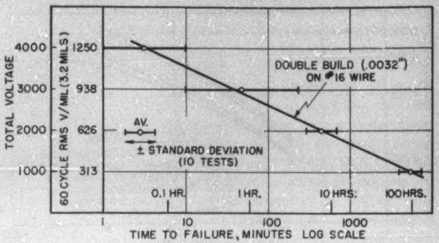


Fig. 8. Corona breakdown of wire enamel between a twisted pair of wires

parable for the same average voltage gradient to the time to breakdown for the nylon and mylar films with metal electrodes. The wire enamel is a phenolic modified polyvinyl formal resin.

An examination of the cross section of one of the enameled wires, shown in Fig. 9, after breakdown, shows that the average erosion is about 30 per cent of the thickness. At points it must exceed this to a depth sufficient enough that the dielectric strength of the remaining film is exceeded. A careful examination of the dark ring, which is the enamel illustrated in Fig. 9, shows that erosion is a maximum at two points (indicated by the arrows) away from the point of contact of the two round wires (indicated by the line). This is because the air gap thickness, which increases from zero at the point of contact, reaches an optimum value for breakdown at these other points, according to Paschen's curve of air breakdown voltage as a function of the product of gap spacing and pressure. Since the air breakdown voltage increases steeply for very small gaps, less erosion is noted very close to the point of contact to the cross section of the other wire (not shown). This photograph was made by casting the wire in a low melting alloy, cutting the casting, and (roughly) polishing it. The vertical streaks are polishing marks.

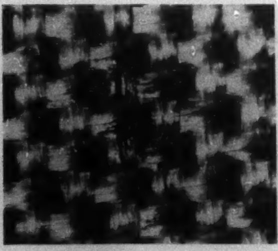


Fig. 7. Corona pit in glass melamine laminate

Voltage Endurance and Corona Intensity

If the decreasing dielectric strength with time of voltage application is due to corona, it can be expected that the time to failure would be inversely related to the corona intensity. The approximate parallel between the time to failure and the reciprocal corona intensity is shown in Fig. 10 for nylon film. The corona intensity measured here is proportional to the voltage amplitude and the number per second of corona discharge pulses. It was obtained using a wide-band amplifier with an output meter which had a long time constant. The proportionality to both voltage and number of pulses per second was determined by using a calibrating pulse generator.

An important factor affecting the corona discharge repetition rate is the frequency of alternation of the voltage. With direct voltage, a corona discharge streamer in series with a dielectric surface charges the surface to a voltage which prevents further discharge streamers from occurring until the applied voltage increases further, decreases, or reverses or until the surface charge leaks off by surface or volume conductivity. Thus, with a steady direct voltage, corona discharges appear very intermittently. It was determined that in the case of a solid nonporous sheet of dielectric, electrode edge corona in air with direct voltage was less by a factor as much as 10^5 than with 60-cycle voltage of equal crest value.

In fact, the d-c corona discharges came at such a low rate that the output intensity meter time constant was not long enough to record a steady reading. Some d-c corona measurements attempted with a porous insulation (an unimpregnated sheet of paper pressboard) showed higher relative intensity than with electrode edge corona. However, the intensity was still lower by a factor of the order of 100-to-1,000 than with 60-cycle a-c corona on the same material.

If the corona discharge rate is limited by surface charges, then increasing the frequency of voltage alternation will increase the repetition rate of discharges and correspondingly the rate of attack on the insulation. British authors[9-11] have done a thorough investigation of corona discharges with internal voids, especially in polyethylene cable. They find that the life of cables which have discharges is inversely proportional to the frequency of the applied voltage. It is possible that some of the observed decreasing dielectric strength of insulation with frequency is due to corona. How big a factor the corona is compared to dielectric heating in higher frequency breakdown needs further study.

Effect of Dielectric Thickness on Time to Failure

Tests of corona breakdown in air are much easier with thinner films than with thick, because edge corona on thick sheets of insulation spreads out to such an extent that surface flash-around becomes a serious problem. Consequently, a majority of the corona breakdown tests described in this paper have been done with thinner dielectrics. Tests at lower aver-

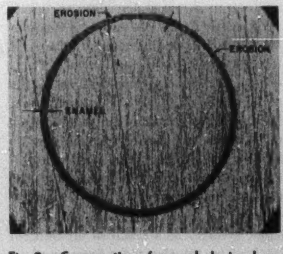


Fig. 9. Cross section of enameled wire showing corona erosion

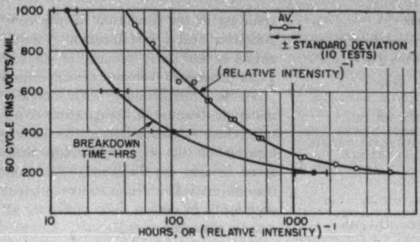


Fig. 10. Corona intensity and time to breakdown, 5-mil nylon

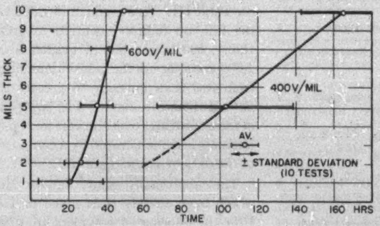


Fig. 11. Voltage endurance, nylon film

age voltage gradients can be done without much difficulty, but they require long times to complete. Some of these are under way at present.

In Fig. 11, tests on nylon film show that at the same voltage gradient the thickness has less effect on the time to breakdown at the higher gradients. The relatively small variation of time to breakdown with thickness for the same voltage gradient is also shown in Fig. 5 where 2-mil mylar is compared at average gradients of 600 volts per mil and above.

Tests have been made on polyethylene at thicknesses from 4 mils to 30 mils, each at the same gradient of 600 volts per mil. Over this wider range of thickness, the time to breakdown did not vary appreciably. Tests are now in progress at lower average gradients on this material.

A test on a 125-mil polyester-styrene polymer sheet at 240 volts per mil produced failure in 1,007 hours. It can be estimated from data on thinner sheets that a 5-mil film of this material would have failed in about 600 hours at this same gradient. This comparison shows even less dependence of the breakdown time on the thickness for the same applied gradient than is indicated by Fig. 11. This polymer sheet after failure is shown in Fig. 12. Deep etching by the corona radiates away from the electrode edge with little evidence of carbonization except at the breakdown hole.

Dependence of Voltage Endurance on Materials and the Mechanism of Breakdown

To compare the voltage endurance of different materials, it is necessary that they be tested at the same voltage gradient and preferably also the same thickness, although the latter factor appears to be less critical. An examination of Table I will show that all of the plastic materials have about the same order of voltage endurance with the exception of the fluorinated polymers, Kel F and Teflon. At least two of the resin-impregnated mica papers appear to be much more resistant to corona than the resins alone, and the mica splittings are very superior.

In the case of the mica papers, it seems that the mica flakes protect the resin beneath from the corona bombardment and give in some cases a large measure of improvement in the corona resistance. Another example of the protective effect of an inorganic filler on the corona resistance is found in silicone rubber, some data for which are shown in Table II. Comparative data are also given for an unfilled resin coated, by dipping, onto a glass cloth. It can be seen that the unfilled silicone resin is only a little better if any, than the other organic resins in Table I (with the exception of the fluorinated resins). The glass cloth may give slight protection to the resin, but some tests on the silicone rubbers show very long times to failure. Unfortunately, the thinner samples for which tests are shown here have low average times to failure. A few tests which have lasted very long

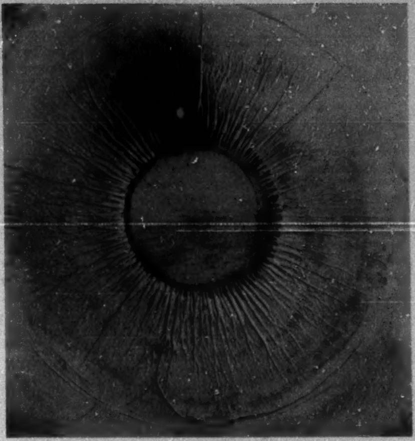


Fig. 12. Corona erosion breakdown of 1/8-inch plastic sheet

Table I. Comparison of Voltage Endurance of Film Materials in Presence of Corona

Material	Thickness, Mils	Volts per Mil	Hours to Breakdown: Average	Hours to Breakdown: Standard Deviation
Polyethylene	4.4	600	53	20
Mylar	8.4	600	27.2	8
Polystyrene	5	600	74	22
Nylon	5	600	36	7
Silicone*	6.5	615	74	46
Polyester-styrene Resin	5	583	33	12
Kel F	5	600	7.1	2.3
Teflon	5	600	5.7	4.3
Mica paper A†	5.2	577	4,100	750
Mica paper B†	5.2	577	109	40
Mica paper C†	6.0	510	1,218	383
Mica splittings	1	1,000	>1,500	no failure
Mica splittings	1	2,000	572	519

* General Electric *81132* on 4-mil glass cloth.

† Experimental 3-sheet laminates vacuum-impregnated with organic solventless thermosetting resin.

Table II. Voltage Endurance, Silicones in Presence of Corona

Silicone Rubbers	Thickness, Mils	Volts per Mil	No. Tests	Time to Failure, Hours	
				Average	Range
No. 1	9	324	5	324	(50, 157, 365, 525, 525)
No. 2	8	400	5		(62, 80, 91, 181, >3,450)
No. 3	7.5	445	5		(0.1, 0.2, 706, 1,192, >3,478)
Glass backed wrapper	11.5	400	5	152	(95, 96.5, 153, 182, 234)
Silicone Resins Treated Glass Cloth*					
Resin A	6.5	615	10	74	(Standard Deviation 46)
Resin B	6.5	615	10	168	(Standard Deviation 88)

* 4-mil glass cloth treated by dipping.

Table III. Voltage Endurance of Mica Splittings* in Presence of Corona

Single Splittings	Thickness, Mils	Volts per Mil	No. Tests	Time to Failure, Hours	
				Average	Range
	1	1,000	6	>1,524	No failures
	1	2,245	8	572	(0.3, 0.3, 99, 389, 769, 818, 1,102>1,485)
	1	2,600	10	800	(0.9, 1.3, 41, 41, 853, 864, 1,817, 2,346, others >2,875)
Stacked Splittings					
Two 1-mil sheets	2	1,500	5	680	(0.9, 455, 670, 1,086, 1,216)
Two 1-mil sheets	2	2,000	5	1.8	(0.2, 0.3, 2.2, 2.6, 3.2)
Five 1-mil sheets (Sheets 2 and 4 separated at center under electrode)	5	900	5	>2,700	(2,712, others >2,800)

* 4-firsts Indian Muscovite Mica.

without failure indicate the high intrinsic corona resistance possessed by this material with the inorganic (silica) filler.

The wide dispersion of the test results obtained with the silicone rubbers and also with the mica papers illustrates the importance of defects in determining the voltage endurance. The voltage endurance appears to be affected by defects in much the same way as the dielectric strength. If there is a point within the area covered by the corona which is weaker in dielectric strength or in resistance to corona, the corona streamers will seek it out and erode through at that point. Filled materials which derive their corona resistance from the filler, such as the silicone rubber and the mica papers, are much more likely to have these weak points since the filler is susceptible to nonuniform dispersion. Furthermore, thread or filaments of other organic material (dust, for example), if present in the material, can easily form paths of low resistance to corona nearly through the material.

Thicker sheets of the silicone rubber are likely to be uniformly better, since it is less likely that a defect or impurity would be large enough to penetrate a large portion of the thickness.

The mechanism of the corona attack is believed to be largely an electron bombardment of the material. Energetic electrons present in the corona streamers impinge on the surface, exciting the molecules sufficiently to produce rupture of chemical bonds and molecular fragments, which often recombine to low molecular weight volatile products.

The process can be considered in many respects analogous to the disruption of molecules by electrons in the ionization chamber of a mass spectrograph. Ozone attack can also occur, but is probably secondary, since tests on plastics in inert gases (Argon and Neon) by Brodhun and Perkins[12,13] produced failures in about the same time as in air.

Organic materials are susceptible to decomposition by energetic electrons to low molecular weight volatile products, but inorganic materials are less so. The structure of inorganic materials can be slowly disrupted by energetic electrons, resulting in the formation of new compounds or local recrystallization. Volatilization is less likely because of the low vapor pressure of inorganic decomposition products. Furthermore, with inorganic materials, because of their regular structure, there is greater chance for recombination of the original structure after atoms or ions are displaced. For these reasons mica has demonstrated a great superiority in corona resistance over all organic materials tested.

Mica, however, has weak points, as shown in a number of tests at high gradients by the wide dispersion of test results; see Table III. At a gradient of 1,000 volts per mil, no failures occurred in greater than 1,500 hours, but at gradients over 2,000 volts per mil the failure times ranged from less than an hour to greater than several thousand hours. The superiority of mica splittings over the best mica paper laminate is demonstrated by a comparsion of the data in Tables I and III.

The abnormally short failure times which have already been noted in Table I for the fluorinated polymers are understandable if it is pointed out that these fluorinated compounds are subject to decomposition by low energy electrons of 1 to 2 electron volts by virtue of the large electron affinity of the fluorine. The release of a fluorine ion, which can initiate the decomposition, can be considered as the sum of two reactions for which the energy relations are approximately known:

$$\begin{array}{l} RF \rightarrow R^{\cdot} + F - 4.54 \text{ electron volts} \\ e + F \rightarrow F^{-} + 3.58 \text{ electron volts} \\ \hline e + RF \rightarrow R^{\cdot} + F^{-} - 0.97 \text{ electron volt} \end{array}$$

In the foregoing equations R represents the remainder of the carbon-fluorinemolecule. The carbon-fluorine bond energy given was that for a partially fluorinated compound. It may be somewhat higher for a completely fluorinated compound.

No reaction requiring such low electron energies can readily be postulated for other nonhalogenated organic polymers, but since a reaction is possible in the fluorine polymers it can be understood why the decomposition of such polymers

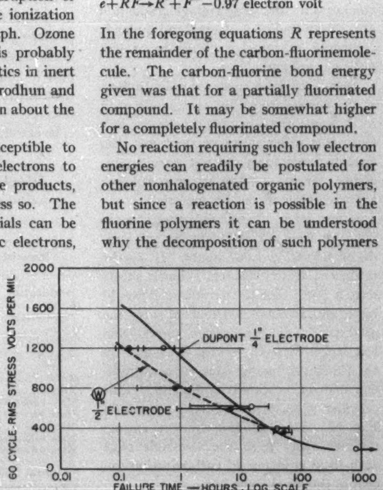


Fig. 13 (right). Voltage endurance of 2x1.75 mils of Teflon film

would proceed more rapidly in the presence of corona. At any given voltage gradient, corona streamers would contain more electrons of sufficient energy to decompose these fluorinated polymers.

Tests which were made on Teflon demonstrate another effect of defects on the voltage endurance. To check the voltage endurance of Teflon, parallel tests were run in the DuPont Laboratory by Mr. Brodhun and in the Westinghouse Research Laboratory. The same sheets of plastic were tested in both laboratories, but 1/4-inch-diameter electrodes were used in the DuPont Laboratory and 1/2-inch-diameter electrodes in the Westinghouse Laboratory. The results shown in Fig. 13 were obtained. Use of a larger electrode produced shorter average breakdown times. Such behavior can be expected where data show a dispersion due to weak points. Larger area electrodes include more of the weaker points and give average results closer to the minimum of tests with smaller electrodes.

It should not be construed that Teflon is unique in this respect. In Table I are given the standard deviations of test results on other plastic films and from these it can be predicted that larger electrodes would have given shorter breakdown times and how much shorter.

Summary

The effect of corona on the long-time dielectric strength of insulation is emphasized in this paper, and various factors influencing the breakdown of insulation with corona are discussed. Data on the corona threshold voltages on dielectric surfaces at electrode edges show that the threshold voltage varies for a wide variety of dielectrics only as the ratio of dielectric thickness to dielectric constant.

The time to failure with corona varies inversely as the corona intensity, which is very low with direct current and increases with the frequency of applied voltage. It is shown that the dielectric strength in the presence of corona begins decreasing at the first cycle of alternating voltage and appears to level off only at the corona starting voltage.

The voltage endurance with corona does not vary much with dielectric thickness of unfilled resins for the same applied voltage gradient.

It has been noted that the time to failure with corona is strongly affected by weak points or defects in the insulation, which produce early failures. This is especially true with inorganic filled materials which derive their corona resistance from the filler, notably impregnated-mica papers and silicone rubber. Fluorinated polymers have been shown to erode more rapidly with corona due to the low-electron energy required to decompose them. Other organic resins tested all appear to have the same order of magnitude of resistance to corona.

References

1. Dielectric Breakdown of Solids (book), **S. Whitehead.** Clardenden Press, Oxford, England, chap. IV, "Breakdown Caused by Discharges," 1951. This reference reviews the literature in this subject.

2. Dielectric Phenomena in High Voltage Engineering, **F. W. Peek, Jr.** McGraw-Hill Book Company, Inc., New York, N. Y., 1929, p. 245.

3. Dielectric Properties of Fibrous Insulation as Affected by Repeated Voltage Application, **F. M. Clark.** *AIEE Transactions,* vol. 44, 1925, pp. 193–210.

4. Breakdown Curve for Solid Insulation, **V. M. Montsinger.** *AIEE Transactions (Electrical Engineering),* vol. 54, Dec. 1935, pp. 1300–01.

5. Dielectric Strength of Transformer Insulation, **P. L. Bellaschi, W. L. Teague.** *AIEE Transactions (Electrical Engineering),* vol. 56, Jan. 1937, pp. 164–71.

6. Impulse Dielectric Strength Characteristics of Liquid-Impregnated Pressboard, **T. W. Dakin, C. N. Works.** *AIEE Transactions,* vol. 71, pt. I, 1952, pp. 321–25.

7. Alternating and Direct Voltage Endurance Studies on Mica Insulation for Electric Machinery, **Graham Lee Moses.** *AIEE Transactions,* vol. 70, pt. I, 1951, pp. 763–69.

8. Dielectric Fatigue of Thermoset Laminates, **N. A. Snow, C. A. Belsterling.** *Modern Plastics,* New York, N. Y., vol. 27, July 1950, p. 93.

9. The Effect of Electric Stress on the Life of Cables Incorporating a Polythene Dielectric, **P. R. Howard.** *Journal,* Institution of Electrical Engineers, London, England, vol. 98, pt. II, 1951, pp. 365–70.

10. Internal Discharges in Dielectrics, Their Observation and Analysis, **A. E. W. Austen, W. Hackett.** *Journal,* Institution of Electrical Engineers, London, England, vol. 91, pt. I, 1944, p. 298.

11. The Voltage Characteristics of Polythylene Cables, **R. Davis, A. E. W. Austen, Willis Jackson.** *Journal,* Institution of Electrical Engineers, London, England, vol. 94, pt. III, 1947, p. 154.

12. Teflon, Polytetrafluoroethylene Resin as an Electrical Insulator, **J. R. Perkins, C. G. Brodhun.** *Annual Report,* 1949 Conference on Electrical Insulation, National Research Council, Washington, D. C., 1950.

13. Dielectric Breakdown of Polymers under Prolonged Stress, **Carl Brodhun.** *Physical Review,* New York, N. Y., vol. 86, 1952, p. 653A.

◆

Discussion

G. M. L. Sommerman (Battelle Memorial Institute, Columbus, Ohio): The curves of constant-stress life versus thickness in Fig. 11 are very interesting, in that they show an increase in life with increasing thickness. In ionization-free short-time breakdown tests, other investigations have yielded data which can be transformed to constant-stress life, $T=kd^n$, where d is the sample thickness and n is less than, or equal to, zero, depending on whether or not thermal effects are present. The fact that n is greater than zero in the authors' tests means that substantially more time is required for ionization of approximately the same intensity to cut through the thicker samples to the point where complete failure finally occurs. The last stage of failure may be disruptive or thermal in character, depending on the electric conductivity of the sample, and, once started, probably proceeds very rapidly.

It is probable that thermal effects do not influence significantly the insulation life measured in these tests. This is because the tests were made at room temperature, and, in many cases, on relatively thin samples. The conventional way of checking for thermal effects is to make other tests at higher temperatures. Another way, applicable in these tests, is to make other tests at higher air pressures. These tests may not be conclusive unless the pressure is raised enough to eliminate corona altogether.

It would be of interest to know if the corona-detection equipment was used in the life tests on liquid-immersed samples included in Fig. 5. If so, was corona detected?

It would also be of interest to know if the samples were conditioned at a definite humidity prior to and during the tests. This may not be very important for many of the materials tested, but for nylon, which absorbs an appreciable amount of moisture at even moderate humidities, there may be a significant effect.

T. W. Liao (General Electric Company, Pittsfield, Mass.): The results obtained by the authors are very interesting and informative. The emphasis of the paper is more on the tests in air than in oil. Accordingly, the reasons given for the short-time failures in oil are not as extensive and positive as the rest of the discussions. The following results, which were obtained at the General Electric High Voltage Laboratory in 1950, might give insight into the breakdown process. They specifically support one of the authors' explanations, namely: "Even in the absence of bubbles initially—the local breakdown of the oil will produce gas, forming bubbles."

Electric discharges in oil are liable to

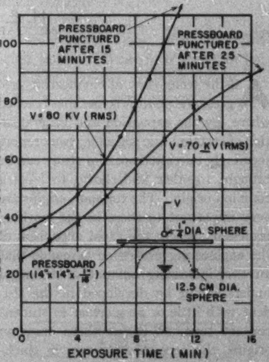


Fig. 14. Increase in relative discharge intensity with time

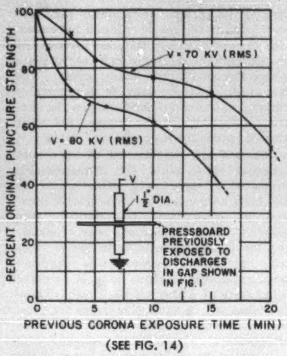


Fig. 15. Per-cent reduction of puncture strength of pressboard previously exposed to discharges

cause damage to organic solid insulation. Once they start to impinge on the insulation, the intensity usually increases rapidly with time, resulting in a progressive deterioration, and finally an ultimate breakdown of the solid insulation. Fig. 14 shows the increase in the relative discharge intensity with time from a small-diameter-sphere gap with an intervening oil-impregnated pressboard barrier for two applied voltages. Gas bubbles were observed and progressive visible damage was noted. The increase in discharge intensity was apparently due to both bubbles and the deterioration of the pressboard. Accordingly, the electric strength of the pressboard decreased with the time of exposure to the discharge. Fig. 15 shows the per-cent reduction in the puncture strength of the same pressboard sheets exposed to the electric discharges of Fig. 14. The puncture strength, determined with two $1^1/_2$-inch-diameter flat electrodes, decreased considerably for the long corona exposure, indicating electrically the damage produced by the discharges. A statement can thus be made from the foregoing results that the intensity of discharges in oil increases rapidly with time, causing progressive weakening of the insulation, and consequently resulting in a short time to failure.

T. W. Dakin, H. M. Philofsky, and **W. C. Divens:** Dr. Sommerman referred to tests presumably described in the literature, which showed a decreasing or constant voltage endurance with increasing sample thickness at a constant voltage stress. This is in contrast to the results described in this paper, and to more recent tests not yet published extending those reported here. It has been found to be consistently the case that at lower stresses of the order of 400 volts per mil or less that the voltage endurance increases with thickness, and more steeply the lower the voltage stress. At higher stresses, the variation with thickness tends to decrease. The tests referred to by Dr. Sommerman could possibly have been tests in which internal dielectric heating occurred. This would have resulted in a higher internal temperature with thicker samples and the sort of dependence quoted by him.

We have used our corona-detection apparatus for detecting discharges with liquid-immersed solid dielectrics. The intensity of discharges at the threshold voltage is extremely low, but is observed to increase rapidly with time at voltages somewhat above the threshold. This is in agreement with the results reported by Dr. Liao. The increasing intensity with time is believed to be associated with the increasing size and number of gas bubbles. The gas is formed by decomposition of the liquid, and is, in the case of mineral oil, very likely largely hydrogen. Some gas may also come from decomposition of the solid. The majority of our work in detection of discharges in liquids has been on thin capacitor dielectrics where such discharges lead to failure in time when they continue to occur. The phenomenon is general, however, wherever the dielectric strength of the liquid is exceeded locally, and is not restricted to any particular system.

The tests described by Dr. Liao confirm very clearly the permanent nature of the destruction caused by discharges at electrode edges in alternating voltage tests. This, of course, is what produces the decreasing dielectric strength of pressboard with time as discussed in reference 6 of the paper. Most of the tests described in the paper were run at an ambient humidity of about 40 per cent (relative), but no attempt was made to control it.

Application of Network Theory to the Analysis of Rotating Machinery

Part I—Synchronous and Asynchronous Machines

H. E. KOENIG
ASSOCIATE MEMBER AIEE

As the technology of engineering progresses a point is sometimes reached where, by the expansion and generalization of certain basic concepts, two or more seemingly unrelated areas of study can be brought together and shown to have a common origin. The common origins are very important to those in the teaching profession, where the rapid expansion of the engineering profession demands that an ever-increasing amount of technical understanding be developed in the student with little or no increase in student contact time. Without these common origins, the necessary increase in teaching efficiency would be difficult to attain. One of the purposes of this, the companion paper,[1] and a conference paper, "General Network Theory in Terms of Matrix Algebra" by M. B. Reed, is to show that by expanding our general network theory and concepts as we now think of them, and by taking a slightly different and more general viewpoint toward the rotating machinery problem, these two areas of technology can be built on a common and rigorous foundation which results in a clearer and broader understanding of both areas, and at the same time will perhaps provide for a more efficient teaching program.

For those engaged in industrial activity, it is believed that the type of analysis presented here has practical value in that it presents a general method whereby all types of rotating machinery can be analyzed along with their associated networks to obtain both steady-state and transient performances. The rotating-field concepts and the steady-state characteristics, as we usually think of them, appear as a particular solution to the more general problem. Thus, the thesis here is not a question of abandoning all the present concepts and techniques used in machine analysis in favor of new ones, but of expanding the present view-points to develop a method for solving the more complex problem.

Paper 54-87, recommended by the AIEE Basic Sciences and Rotating Machinery Committees and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 19, 1953; made available for printing November 25, 1953.

H. E. Koenig is with the University of Illinois, Urbana, Ill.

Acknowledgment is made to the Graduate College of the University of Illinois for permission to utilize material in a thesis, "A Circuit Approach to the Analysis of Electrical Machinery," presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical Engineering. The author also acknowledges the encouragement and material assistance received from the Reliance Electric and Engineering Company and Drs. G. H. Fett and M. B. Reed, Electrical Engineering Department, University of Illinois.

If the general network theory presented in M. B. Reed's paper is used as a starting point, the problem of analyzing rotating machinery appears simply as an electric network problem in which there is relative motion between some or all of the elements of the network. Thus, perhaps it would be well to refer to the rotating machine as a "dynamic network." To extend the general theory of static networks to the dynamic case, it is only necessary to develop methods for handling the variable coefficients that appear as a result of the relative motion between the rotor and stator.

It is the purpose of this paper to extend the static network theory presented in M. B. Reed's paper to the dynamic case, and to outline a general and rigorous method for analyzing the rotating machine and its associated network by a single basic technique applicable to all types of machines. The synchronous and induction machines are used as examples to demonstrate the general procedure for noncommutating machines. The problem of the commutating machine is considered in reference 1.

The treatment given in this paper is not intended to be an exhaustive one, but rather an introduction to the author's viewpoints on the subject. It is hoped that the viewpoints presented here will help to stimulate new interest in the rotating machinery problems. The idea that the rotating machine can be considered as a network with time-varying coefficients is not new, and many of the results arrived at in this paper have been arrived at by previous authors using many varied methods of analysis.[2,3,4] However, the analysis as presented in this paper avoids the undue complexities of "primitive network" concepts and tensor analysis present in much of the previous work in this area. To repeat, it is believed that the unity and fundamental characteristics of the method presented here will lead to a greatly enhanced understanding of rotating machinery problems and to a marked increase in the complexity of the problems that can be solved.

The machine constants used in these analyses can be evaluated by familiar and conventional methods if the proper interpretation is used.

Nomenclature

Nomenclature not confined in the paper or which is not obvious from the network diagrams is defined as follows:

$P=\frac{d}{dt}$
r_p = phase resistance of primary (armature) windings
r_s = phase resistance of secondary windings
L_d = maximum inductance of primary (armature) phases
L_q = minimum inductance of primary (armature) phases
$L_0=(L_d+L_q)/2$
$L_v=(L_d-L_q)/2$
M_d = maximum coupling between phases
M_q = minimum coupling between phases
$M_0=(M_d+M_q)/2$
$M_v=(M_d-M_q)/2$
L_s = inductance of secondary phases
M_s = coupling between secondary phases
M_{rs} = maximum coupling between rotor and stator mesh circuits
ϕ = angular position of rotor
δ = angular displacement between primary (armature) phases
Δ = angular displacement between secondary phases.

Assumptions, Reference Systems, and Definitions

To aid in formulating the problem, it is convenient to consider the analysis of the rotating machine in three major parts, as follows:

1. Establish a set of differential equations describing the performance of the machine in terms of circuit parameters.
2. Evaluate the circuit parameters in terms of machine dimensions.
3. Solve the resulting equations to obtain the general transient response or the frequency response.

Only the first part of the problem is considered in this paper. Unless otherwise stated, the analysis that follows is intended to be rigorous and complete for machines satisfying the following idealizations:

1. Both sides of the air gap are to be free from winding slots.
2. All inductances of the machine are independent of current over the operating range.
3. The iron portions of the machine are free of eddy currents.
4. Capacitance between conductors is negligible.
5. All harmonics above the second are negligible.

If only the steady-state solution is desired, the first and last limitation can easily be relaxed. For balanced conditions of operation, the variation in the inductance terms with current can be included by graphical means.

For the sake of simplicity, the analyses are in terms of one pair of poles and all angles are in electrical radians or degrees. Since only the relative motion between the rotor and stator is important, the armature will always be taken as the rotating member. All equations are written in matrix form following the conventions used in M. B. Reed's paper.

Since the inductance coefficients may be functions of time, it is necessary to write inductance voltage expressions as

$$\frac{d}{dt}(M_{jk}i_k)=M_{jk}\frac{d}{dt}i_k+i_k\frac{d}{dt}M_{jk}$$

Magnetic polarization marks are used on all coils of the machine, their location with respect to the terminals being completely arbitrary. However, once they are established they determine uniquely the sign on both the mutual inductance voltage terms, $(d/dt)(M_{jk}i_k)$, in the differential equations, and the sign of the mutual inductance M_{jk}, in the following way:

1. The $(d/dt)(M_{jk}i_k)$ term is entered in the voltage equation with the same sign as the $(d/dt)(L_ji_j)$ term if the jth and kth reference currents are both into (away) from the marked terminals. Under all other conditions the signs are opposite.
2. When the physical orientation of the two coils in question is specified, the magnitude and sign of the mutual inductance M_{jk} can be evaluated by passing a current into the marked terminal of the kth coil and evaluating the integral $\int \mathbf{B}_k \cdot d\mathbf{A}$ over the area of the jth coil. Positive area is on the left of an observer traveling around the winding from the marked to the unmarked terminal; $d\mathbf{A}$ is the outward normal to the area; and $\mathbf{B}_k$ is the vector magnetic-flux density resulting from the current in the kth coil.

The 2-Winding Transformer

To help the uninitiated to become oriented into the techniques of matrix algebra and to develop the concept of the transformation of variable used in this paper, matrix algebra is used to derive the equivalent circuit of the 2-winding transformer, a problem which is familiar and to which the answer is known.

The connection diagram for a 2-winding transformer along with an appropriate set of mesh circuit is shown in Fig. 1. Using standard network theory, the two complex impedance equations for this network are

$$V_1=(r_1+j\omega L_1)I_1-j\omega M_{12}I_2 \tag{1A}$$

$$-V_2=-j\omega M_{12}I_1+(r_2+j\omega L_2)I_2 \tag{1B}$$

or in matrix form

$$\begin{bmatrix} V_1 \\ -V_2 \end{bmatrix}=\begin{bmatrix} Z_1 & -Z_{12} \\ -Z_{12} & Z_2 \end{bmatrix}\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \tag{2}$$

where

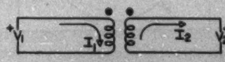


Fig. 1. Connection diagram of a 2-winding transformer

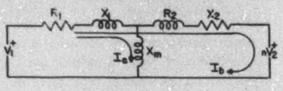


Fig. 2. Equivalent circuit of a 2-winding transformer

$\boldsymbol{Z}_1 = r_1 + j\omega L_1$
$\boldsymbol{Z}_2 = r_2 + j\omega L_2$
$\boldsymbol{Z}_{12} = j\omega M_{12}$

Using the symbolic notation of matrix algebra, equation 2 can be written simply as

$$\mathcal{V} = \mathcal{Z}\mathcal{I} \tag{3}$$

In the usual transformer, L_1I_1 and $M_{12}I_2$ are of the same order of magnitude and nonlinear. If a transformation of variable can be effected which results in coefficients which are the difference of these two quantities, a much higher degree of accuracy will result and the coefficients will be more nearly linear.

Define a new set of voltage and current variables such that

$$\mathcal{V}_c = \mathcal{C}^1\mathcal{V} \text{ and } \mathcal{I}_c = \mathcal{C}^{-1}\mathcal{I} \tag{4}$$

where

$$\mathcal{C} = \begin{bmatrix} 1 & 1 \\ 0 & n \end{bmatrix}$$

$$\mathcal{C}^{-1} = \begin{bmatrix} 1 & -1/n \\ 0 & 1/n \end{bmatrix} \tag{5}$$

n = turns ratio of the transformer, N_1/N_2.

The transformation of variable is formally effected by multiplying equation 3 by $\mathcal{C}^1$ and replacing $\mathcal{I}$ by $\mathcal{C}\mathcal{I}_c$.

$$\mathcal{V}_c = \mathcal{C}^1\mathcal{Z}\mathcal{C}\mathcal{I}_c \tag{6}$$

By forming the matrix products indicated in equations 4 and 6, the detailed expression of the equations of the network in terms of the new variables is

$$\begin{bmatrix} V_1 \\ V_1 - nV_2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{Z}_1 & (\boldsymbol{Z}_1 - n\boldsymbol{Z}_{12}) \\ (\boldsymbol{Z}_1 - n\boldsymbol{Z}_{12}) & (\boldsymbol{Z}_1 - n\boldsymbol{Z}_{12} + n^2\boldsymbol{Z}_2 - n\boldsymbol{Z}_{12}) \end{bmatrix} \times \begin{bmatrix} I_a \\ I_b \end{bmatrix} \tag{7}$$

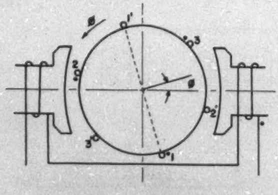


Fig. 3. Map showing physical orientation of machine elements

These two equations can now be represented by the network of Fig. 2 where

$X_1 = \omega(L_1 - nM_{12})$ = leakage reactance of the primary
$X_2 = \omega(n^2L_2 - nM_{12})$ = leakage reactance of the secondary referred to the primary
$R_1 = r_1$
$R_2 = n^2r_2$
$X_m = \omega nM_{12}$ = magnetizing reactance

From empirical considerations, a resistance can be added in parallel with X_m to approximate the core losses.

By studying the symmetry of the impedance matrix and the results presented here, it becomes evident that the same type of transformation of variable can be used to an advantage in the study of the multiwinding transformer. All leakage reactances are proportional to the difference between two inductances, the inductances being defined in terms of flux-density integrals over certain well defined areas. Since there are no definite laws of network synthesis available, it may be difficult to draw a circuit corresponding to the transformed equations of the multiwinding transformer. However, this does not destroy the usefulness of the transformation, for the equivalent circuit diagram only serves to give a pictorial representation of the information already contained in the equations.

Unfortunately there is no apparent direct method available for deriving the transformation used in this and the following examples. Indeed, there is more than one transformation that will accomplish the same objective. Usually the transformations result from a thorough study of the symmetry of the $\mathcal{Z}$ matrix and by a method of cut and try.

Three-Phase Synchronous Machines

Just as in the analysis of any complex network, a good map of the machine is most essential. In addition to showing the connections between the elements of the network, a good map must contain:

1. A complete set of references for all element voltages.
2. A complete set of references for all mesh currents.
3. All information necessary to determine the self and mutual inductances as a function of rotor position.

The maps are greatly simplified if inductance functions are determined from one map and then transferred to the second map by way of the polarization marks discussed in M. B. Reed's paper. Figs. 3 and 4 are examples of the two maps for a 3-phase synchronous machine without damper windings. Although the damper windings could be included, they are omitted in this example for the sake of simplicity.

Having established the necessary maps showing all the variables and their references, the differential equations can now be written for the network. For the present example, the network is sufficiently simple to allow writing the differential equations in terms of mesh-circuit variables directly from inspection of the network. As shown in reference 1, it is sometimes necessary to write the equations first in terms of element variables and then, by means of a matrix transformation of variable, express them in terms of the mesh-circuit variables.

The mesh-circuit differential equations for the network in Fig. 4, as written from inspection of the network, are

$$\left[\begin{array}{c} -v_1 \\ -v_2 \\ -v_3 \\ \hdashline -v_4 \end{array}\right] = \left[\begin{array}{ccc:c} r_1+PL_1 & PM_{12} & PM_{13} & PM_{14} \\ PM_{21} & r_2+PL_2 & PM_{23} & PM_{24} \\ PM_{31} & PM_{32} & r_3+PL_3 & PM_{32} \\ \hdashline PM_{41} & PM_{42} & PM_{43} & r_4+PL_4 \end{array}\right] \left[\begin{array}{c} i_1 \\ i_2 \\ i_3 \\ \hdashline i_4 \end{array}\right] \tag{8}$$

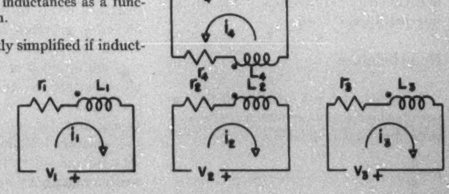
To aid in manipulation, it is convenient to partition the three matrices of equation 8 into submatrices as indicated by the dashed lines. A matrix equation can be partitioned into any conformable set of submatrices.

Equation 8 can now be written in terms of the submatrices as

$$\begin{bmatrix} \mathcal{V}_1 \\ \mathcal{V}_2 \end{bmatrix} = \begin{bmatrix} \mathfrak{D}_{11} & \mathfrak{D}_{12} \\ \mathfrak{D}_{21} & \mathfrak{D}_{22} \end{bmatrix} \begin{bmatrix} \mathcal{I}_1 \\ \mathcal{I}_2 \end{bmatrix} \tag{9}$$

where $\mathfrak{D}_{11}$, $\mathcal{V}_1$, and $\mathcal{I}_1$ are derivative functions, voltages and currents of the rotor (armature) circuits. $\mathfrak{D}_{22}$, $\mathcal{V}_2$ and $\mathcal{I}_2$ are derivative functions, voltages and cur-

Fig. 4 (right). Connection diagram of a synchronous machine



rents of the stator circuits. $\mathfrak{D}_{12}$ is the transpose of $\mathfrak{D}_{21}$ and represents the coupling between the rotor and stator circuits.

The self- and mutual inductance terms appearing in each of the submatrices can be determined as a function of rotor position.[5] The results are

$$\mathfrak{D}_{11}=\begin{bmatrix}(r_p+PL_0) & -PM_0 & -PM_0\\ -PM_0 & (r_p+PL_0) & -PM_0\\ -PM_0 & -PM_0 & (r_p+PL_0)\end{bmatrix}+ P\begin{bmatrix}L_v\cos 2\phi & M_v\cos(2\phi-\delta) & M_v\cos(2\phi-2\delta)\\ M_v\cos(2\phi-\delta) & L_v\cos(2\phi-2\delta) & M_v\cos 2\phi\\ M_v\cos(2\phi-2\delta) & M_v\cos 2\phi & L_v\cos(2\phi-\delta)\end{bmatrix} \quad (10)$$

$$\mathfrak{D}_{12}=PM_{rs}\begin{bmatrix}\cos\ \phi\\ \cos(\phi-\delta)\\ \cos(\phi-2\delta)\end{bmatrix} \quad (11)$$

$$\mathfrak{D}_{22}=[r_4+PL_4] \quad (12)$$

$$\mathfrak{D}_{21}=PM_{rs}[\cos\phi\ \cos(\phi-\delta)\ \cos(\phi-2\delta)] \quad (13)$$

As a general proposition, it is important to note that the above partitioning is a convenient one to use in machine analysis and that the elements of the $\mathfrak{D}_{11}$ matrix (coefficients of the rotor mesh-circuit equations) and the elements of the $\mathfrak{D}_{22}$ matrix (coefficients of the stator mesh-circuit equations) are functions of the angular displacement between the rotor and stator if the stator and rotor, respectively, have irregular surfaces, whereas they are independent of angular displacement if the stator and rotor have regular cylindrical surfaces. The $\mathfrak{D}_{12}$ and $\mathfrak{D}_{21}$ matrices are always functions of the angular displacement between the rotor and stator. Since the above-mentioned coefficients are always continuous periodic functions of the angular displacement, they can be represented by a Fourier series, and as many harmonics, including slot harmonics, entered in the differential equations as required for a satisfactory solution to the particular problem being studied.

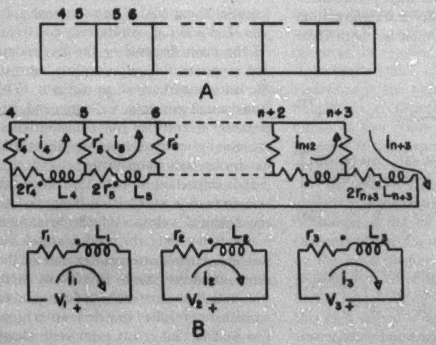


Fig. 5. Induction machine

A. Developed squirrel cage

B. Connection diagram of squirrel cage

Thus far in this paper, only standard network theory has been used in establishing the differential equations for rotating machinery. However, since the resulting differential equations have variable coefficients, it is very difficult to obtain the general (transient) solution to these equations in closed form. The particular (steady-state) solution is usually somewhat easier to obtain, and can be obtained, under certain conditions, for the equations in their present form. The particular techniques to be used in proceeding with the analysis depend entirely upon the solution desired and upon what simplifying assumptions are permissible. It is not the purpose of this paper to investigate the solution to all the special problems in rotating machinery no more than it is the purpose of a paper on network theory to investigate the solution to all possible networks, for this would indeed require volumes. The purpose is rather to outline a general procedure with a few examples included to aid the reader in understanding the general procedure and to show that our present methods of machine analysis are indeed solutions to a more general problem under very special and much simplified conditions.

To accomplish this purpose, consider the problem of obtaining the steady-state solution to the set of differential equations for the synchronous machine as given in equation 8, the steady-state conditions of operation being defined as follows:

1. Constant speed: $\phi=\omega t$.
2. Constant field excitation: r_4= known constant.
3. Continuous periodic armature terminal voltage:

$$\begin{bmatrix}v_1\\ v_2\\ v_3\end{bmatrix}=\begin{bmatrix}|V_1|\sin(\omega t+\theta_1)\\ |V_2|\sin(\omega t+\theta_2)\\ |V_3|\sin(\omega t+\theta_3)\end{bmatrix} \quad (14)$$

Let the analysis be further limited to machines having uniform airgaps. $M_v=L_v=0$. Under these conditions, the steady-state current solution is expected to be of the same form as the terminal voltages, namely

$$\begin{bmatrix}i_1\\ i_2\\ i_3\end{bmatrix}=\begin{bmatrix}|I_1|\sin(\omega t+\psi_1)\\ |I_2|\sin(\omega t+\psi_2)\\ |I_3|\sin(\omega t+\psi_3)\end{bmatrix} \quad (15)$$

When these currents are substituted into the differential equations by forming the matrix products indicated in equation 9, there results

$$\mathfrak{V}_1=r_p\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\mathfrak{I}_1+P\begin{bmatrix}L_0 & -M_0 & -M_0\\ -M_0 & L_0 & -M_0\\ -M_0 & -M_0 & L_0\end{bmatrix}\mathfrak{I}_1+PM_{rs}\begin{bmatrix}\cos\omega t\\ \cos(\omega t-\delta)\\ \cos(\omega t-2\delta)\end{bmatrix}i_4 \quad (16A)$$

$$-v_4=P\frac{M_{rs}}{2}[\,|I_1|\ |I_2|\ |I_3|\,]\begin{bmatrix}\sin\psi_1+\sin(2\omega t+\psi_1)\\ \sin(\psi_2+\delta)+\sin(2\omega t+\psi_2-\delta)\\ \sin(\psi_3+2\delta)+\sin(2\omega t+\psi_3-2\delta)\end{bmatrix}+[r_4+PL_4]i_4 \quad (16B)$$

Note that the matrix product $\mathfrak{D}_{21}\mathfrak{I}_1$ has resulted in two terms, see equation 16 (B), one of which represents an armature-reaction flux-density field which is fixed with respect to the field winding, the other which represents a field rotating at twice synchronous speed with respect to the field windings. The double-frequency voltage terms appearing in equation 16 (B) introduce double-frequency components in the field current, i_4. When i_4 is substituted into equation 16(A), it is clear that third-harmonic components of armature current result, and the assumed form of the armature current was not a correct one. It is also clear that the correct form of the armature currents under the condition of unbalanced terminal voltages would be an infinite series.

If the terminal voltages and currents are balanced

$$|V_1|=|V_2|=|V_3|$$
$$\theta_1=\theta_2+\delta=\theta_3+2\delta$$
$$|I_1|=|I_2|=|I_3|$$
$$\psi_1=\psi_2+\delta=\psi_3+2\delta$$

Under these conditions, the double-frequency voltages appearing in equation 16 (B) sum to zero, leaving the field current i_4 completely independent of the armature currents. The assumed form of the current functions are correct and only the magnitude $|I_1|$ and phase angle ψ_1 of the armature currents remain to be determined. Any one of the three equations represented by matrix equation 16 (A) can be used to determine these constants. In particular, the equation for phase one specifies that

$$-|V_1|\sin(\omega t+\theta_1)=[(r_p+PL_o)-PM_o-PM_o]|I_1|\begin{bmatrix}\sin(\omega t+\psi_1)\\ \sin(\omega t+\psi_1-\delta)\\ \sin(\omega t+\psi_1-2\delta)\end{bmatrix}+i_4PM_{rs}\cos\omega t \qquad (17)$$

Upon following the usual techniques of circuit analysis whereby the sinusoidal functions in linear equations are represented by the imaginary component of the rotating vector $\epsilon^{j\omega t}$, the complex-impedance equivalent of equation 17 becomes

$$-V_1=(r_p+j\omega(L_o-M_o\epsilon^{-j\delta}-M_c\epsilon^{-j2\delta}))I_1-\omega M_{rs}i_4 \qquad (18A)$$

or

$$-V_1=(r_p+j\omega(L_o+M_o))I_1-\omega M_{rs}i_4 \qquad (18B)$$

where V_1 and I_1 are complex numbers.

It can be shown that the inductance of one phase L_o is approximately equal to twice the magnitude of the coupling between phases M_o. If leakage reactance per phase of the machine is defined as

$$\omega L_l=\omega(L_o-2M_o) \qquad (19)$$

then equation 18(B) can be written as

$$-V_1=(r_p+j\omega(L_l+3M_o))I_1-\omega M_{rs}i_4 \qquad (20A)$$

or

$$-V_1=(r_p+jx_s)I_1-E_g \qquad (20B)$$

which is the usual equation for the balanced operation of a synchronous machine with a uniform airgap. The leakage reactance is clearly defined in equation 19 and includes all the various components of leakage reactance appearing in the literature.

If the saturation effects are included, the inductance coefficients in equation 10 become functions of current. The derivative of such inductance-current products takes on the form

$$\frac{d}{dt}L(i)i(t)=L(i)\frac{\partial}{\partial t}i(t)+i(t)\frac{\partial}{\partial i}L(i)\frac{\partial}{\partial t}i(t) \qquad (21)$$

Since the field current in equation 17, i_4, is independent of time, the nonlinearity introduced by saturation of the field poles and frame can be included in the analysis by the familiar graphical methods employing the saturation curve of the machine. However, in the first group of terms on the right side of equation 17, the coefficients are multiplied by time-varying currents which give rise to an additional component of voltage, as indicated in equation 21; a point somewhat obscure in the usual analysis of synchronous machines. In practice, this added component of voltage is approximated by using dynamic saturation curves for the armature iron.

If either the transient or steady-state solution of the salient-pole synchronous machine is required, it is necessary to return to the original form of the differential equations, as given in equation 1, where inspection of the coefficients of the equations reveals a considerable amount of symmetry. That is, the coefficients of the $\mathfrak{D}_{11}$ matrix are all simply related. The symmetry among the coefficients is a direct result of the uniform distribution of equal phase belts over the surface of the armature, and represents a property peculiar to the differential equations of practically every type of rotating machine. It is this symmetry which leads to the important and most useful technique known as a transformation of variable, whereby a new set of current and voltages are defined in terms of the old currents and voltages, and the equations solved in terms of the new variables. Through the inverse transformation, the original currents and voltages can be determined from the new or transformed currents and voltages. When matrix algebra is used, these transformations can be clearly and efficiently effected

If, in equation 8, a transformation of variable can be effected which will diagonalize the $\mathfrak{D}_{11}$ matrix, then the new or transformed rotor currents are all independent. At the same time, it would be desirable to remove the variable coefficients in both the $\mathfrak{D}_{11}$ and the $\mathfrak{D}_{12}$ matrices. The 2-reaction theory introduced by Blondell is in reality a transformation of variable that accomplishes both these desired results.

Let

$$\mathcal{I}_b=\mathcal{B}\mathcal{I}_1 \text{ and } \mathcal{V}_b=\mathcal{B}\mathcal{V}_1 \qquad (22)$$

where

$$\mathcal{B}=\frac{2}{3}\begin{bmatrix}\cos\phi & \cos(\phi-\delta) & \cos(\phi-2\delta)\\ -\sin\phi & -\sin(\phi-\delta) & -\sin(\phi-2\delta)\\ 1/2 & 1/2 & 1/2\end{bmatrix} \qquad (23)$$

The inverse of $\mathcal{B}$ is

$$\mathcal{B}^{-1}=\begin{bmatrix}\cos\phi & -\sin\phi & 1\\ \cos(\phi-\delta) & -\sin(\phi-\delta) & 1\\ \cos(\phi-2\delta) & -\sin(\phi-2\delta) & 1\end{bmatrix} \qquad (24)$$

and the original currents and voltages are related to the transformed currents and voltages by

$$\mathcal{I}_1=\mathcal{B}^{-1}\mathcal{I}_b \text{ and } \mathcal{V}_1=\mathcal{B}^{-1}\mathcal{V}_b \qquad (25)$$

Upon multiplying the top row of equation 9 by $\mathcal{B}$ and replacing $\mathcal{I}_1$ by $\mathcal{B}^{-1}\mathcal{I}_b$ there results,

$$\begin{bmatrix}\mathcal{V}_b\\ \mathcal{V}_2\end{bmatrix}=\begin{bmatrix}\mathfrak{D}_{bb}\mathfrak{D}_{b2}\\ \mathfrak{D}_{2b}\mathfrak{D}_{22}\end{bmatrix}\begin{bmatrix}\mathcal{I}_b\\ \mathcal{I}_2\end{bmatrix} \qquad (26)$$

where

$$\begin{aligned}\mathfrak{D}_{bb}&=\mathcal{B}\mathfrak{D}_{11}\mathcal{B}^{-1}\\ \mathfrak{D}_{b2}&=\mathcal{B}\mathfrak{D}_{12}\\ \mathfrak{D}_{2b}&=\mathfrak{D}_{21}\mathcal{B}^{-1}\end{aligned} \qquad (27)$$

The detailed expression for the matrix equation 26 as obtained by forming the matrix products indicated in equation 27 is

$$\begin{bmatrix}v_b\\ v_q\\ v_o\\ \hline -v_4\end{bmatrix}=\left[\begin{array}{ccc|c}(r_p+L_dP) & L_q\dot{\phi} & 0 & M_{rs}P\\ -L_d\dot{\phi} & (r_p+L_qP) & 0 & M_{rs}\dot{\phi}\\ 0 & 0 & (r_p+L_oP) & 0\\ \hline \frac{3}{2}M_{rs}P & 0 & 0 & (r_4+L_4P)\end{array}\right]\times\begin{bmatrix}i_d\\ i\\ i_o\\ \hline i_4\end{bmatrix} \qquad (28)$$

where

$$L_d=\left(L_o+M_o+\frac{3}{2}L_v\right)$$

$$L_q=\left(L_o+M_o-\frac{3}{2}L_v\right)$$

$$L_o=\left(\frac{L_o}{2}-M_o\right)$$

This transformation of variable has resulted in a set of constant coefficient differential equations. The rotational voltages, $L_d\dot{\phi}$ and $L_q\dot{\phi}$, are the only two terms off the main diagonal of the $\mathfrak{D}_{bb}$ matrix and represent the only coupling between the transformed armature currents. The transformed currents, voltages, and impedances defined by this transformation are usually referred to as the direct and quadrature axis components.

It is important to note that since the transformation of variable is applied to the terminal voltages of the armature, it effects not only the coefficients of the machine, but all external elements in the three armature mesh circuits as well. Also note that since the transformation of variable has resulted in a derivative function matrix that is not symmetric about the main diagonal, the transformed equations cannot be represented by a network of bilateral circuit elements. It is shown in M. B. Reed's paper that there is a certain class of transformations that always results in a symmetric transformed matrix.

Equation 28 can be solved under the most general conditions of operation to obtain the transient or frequency response if the impedances in the 3-mesh circuits of the armature are equal. The steady-state solution of these equations for a set of balanced 3-phase terminal voltages reduces quickly and easily to the relatively simple steady-state 2-reaction theory, if the general procedure of assuming the form of the current solution previously outlined is followed. These equations of course, can also be used to obtain the performance of the uniform-air-gap machine previously discussed, and so represents a more general procedure.

Asynchronous Machines

If the primary or armature of the induction machine is taken as the rotating member, the only difference between the synchronous and asynchronous machines is the number of primary and secondary circuits and the manner in which they are excited. For example, the 3-phase squirrel-cage induction machine has three primary circuits, which are excited in the same manner as those of the synchronous machine. However, the secondary has as many phases or mesh circuits as there are bars, none of which are excited from an external voltage source.

Before the differential equations for the squirrel-cage induction machine can be written, it is necessary to represent the electrical properties of the squirrel cage by an appropriate set of lumped circuit parameters. A diagram of the developed squirrel cage is shown in Fig. 5(A). Since inductance is only defined for closed circuits, it is not possible to define and obtain an inductance coefficients for each bar and end-ring section of the cage. Instead, it is necessary to define the circuit parameters for the secondary in terms of closed paths. If the end ring on one side of the cage is considered to be momentarily opened between successive bars, as shown in Fig. 5(A), then between each pair of terminals, 4–5, 5–6, etc., there is a definite obtainable value of inductance and resistance.

The map showing the connections between the set of circuit parameters describing the electrical properties of the squirrel cage, along with the primary connections, then becomes that shown in Fig. 5(B), where $L_4, L_5 \ldots L_{n+3}$ represent the inductance of circuit 4–5, 5–6 . . . , $(n+3)$–4, respectively. The resistance of one end-ring connection between bars is represented by $r_4, r_5, \ldots, r_{n+3}$. The resistances of the bars are represented by $r_4', r_5', \ldots, r_{n+3}$. Closing the openings in the end rings of Fig. 5(A) is equivalent to short-circuiting points 4,5, . . . , $(n+3)$ in Fig. 5(B). Using the criterion established in M. B. Reed's paper, a satisfactory and convenient set of n mesh-circuit currents for the network are chosen, as indicated in Fig. 5(B).

The matrix equation for the induction machine is the same as that of the synchronous machine given in equation 9. By writing the differential equations for the network of Fig. 5(B), the elements of the $\mathfrak{D}_{11}$ matrix are found to be the same as those given by equation 10, with L_v and M_v set equal to zero. For a squirrel cage rotor with evenly spaced bars, the elements of the $\mathfrak{D}_{12}$ and $\mathfrak{D}_{21}$ matrices are

$$\mathfrak{D}_{12}=\mathfrak{D}_{21}{}^1=PM_{rs}\begin{bmatrix}\cos\phi & \cos(\phi-\Delta) & \ldots\cos[\phi-(n-1)\Delta]\\ \cos(\phi-\delta) & \cos(\phi-\delta-\Delta) & \ldots\cos[\phi-\delta-(n-1)\Delta]\\ \cos(\phi-2\delta) & \cos(\phi-2\delta-\Delta) & \ldots\cos[\phi-2\delta-(n-1)\Delta]\end{bmatrix} \tag{29}$$

where δ and Δ represent the angular displacement between the primary and secondary mesh circuits, respectively, and M_{rs} represents the magnitude of the maximum coupling between a rotor and the stator mesh circuits.

The elements of the $\mathfrak{D}_{22}$ matrix are similarly found to be

$$\begin{bmatrix}R_4+PL_4 & -r_5'+PM_{45} & PM_{46} & \ldots & -r_4'+PM_{4,\,n+3}\\ -r_5'+PM_{45} & R_5+PL_5 & r_6'+PM_{56}\ldots & & PM_{5,\,n+3}\\ PM_{46} & -r_6'+PM_{56} & R_6+PL_6 & \ldots & PM_{6,\,n+3}\\ \vdots & \vdots & \vdots & & \vdots\\ -r_4'+PM_{4,\,n+3} & PM_{5,\,n+3} & PM_{6,\,n+3}\ldots & & R_{n+3}+PL_{n+3}\end{bmatrix} \tag{30}$$

where $R_j=r_j'+r_{j+1}'+2r_j=$resistance of the jth mesh circuit, and M_{jk} represents the magnetic coupling between the jth and kth mesh circuit.

If the armature surface is assumed to be free of winding slots, the elements of the $\mathfrak{D}_{22}$ matrix are independent of rotor position. If it were desirable to study the effects of the winding slots on the steady-state performance of the machine, the variations in the primary and secondary inductance coefficients could be introduced at this point. However, for the purposes of this paper all slot effects are neglected.

If the squirrel cage of the machine has uniformly spaced bars all having the same resistance, it can be shown that all the elements of the $\mathfrak{D}_{22}$ matrix which appear on the main diagonal and all terms appearing on each of the diagonals parallel to the main diagonal, are equal. With this high degree of symmetry among the elements of the network coefficients, it is relatively easy to solve these equations in their present form to determine the steady-state solution for any set of primary terminal voltages, balanced or unbalanced. For convenience, let the matrix equation 9 be rewritten as

$$\mathcal{E}_1=\mathfrak{D}_{11}\mathcal{I}_1+\mathfrak{D}_{12}\mathcal{I}_2 \tag{31A}$$

$$\mathcal{E}_2=\mathfrak{D}_{21}\mathcal{I}_1+\mathfrak{D}_{22}\mathcal{I}_2 \tag{31B}$$

Under steady-state conditions of operation the driving voltages applied to the primary are the same as those given in equation 14. Following the same technique used in the discussion of the synchronous machine, the primary currents are expected to be of the same form as those given in equation 17. The primary currents given in equation 17 can be substituted into equation 31(B) and the secondary currents determined. If the matrix product $\mathfrak{D}_{21}\mathcal{I}_1$ is formed, the usual forward and backward components of revolving fields are obtained and the complete steady-state analysis can be obtained for this unbalanced condition of operation. For balanced conditions of operation, the usual **T**-equivalent circuit is obtained with the leakage reactances in the equivalent circuit defined in terms of the difference in the flux-density integrals over certain well defined areas. Certainly the present methods used to approximate the various components of inductance thus obtained are quite applicable.

If it is necessary to study the transient performance of the induction machine, it is desirable to make a transformation of variable to reduce the complexity of the equations and remove the variable coefficients in the equations. How such a transformation might be determined is shown next.

Let $\mathfrak{B}_t$ and $\mathfrak{B}_e$ be transformation matrices such that

$$\mathcal{I}_b=\mathfrak{B}_t\mathcal{I}_1 \text{ and } \mathcal{E}_b=\mathfrak{B}_e\mathcal{E}_1 \tag{32}$$

and let $\mathfrak{C}_t$ and $\mathfrak{C}_e$ be transformation matrices such that

$$\mathcal{I}_c=\mathfrak{C}_t\mathcal{I}_2 \text{ and } \mathcal{E}_c=\mathfrak{C}_e\mathcal{E}_2 \tag{33}$$

Applying these transformations to equation 31 gives

$$\begin{bmatrix} \mathcal{E}_b \\ \mathcal{E}_c \end{bmatrix} = \begin{bmatrix} \mathfrak{D}_{bb} \mathfrak{D}_{bc} \\ \mathfrak{D}_{cb} \mathfrak{D}_{cc} \end{bmatrix} \begin{bmatrix} \mathcal{I}_b \\ \mathcal{I}_c \end{bmatrix} \tag{34}$$

where

$$\mathfrak{D}_{bb} = \mathfrak{B}_e \mathfrak{D}_{11} \mathfrak{B}_t^{-1} \tag{34A}$$
$$\mathfrak{D}_{bc} = \mathfrak{B}_e \mathfrak{D}_{12} \mathfrak{C}_t^{-1} \tag{34B}$$
$$\mathfrak{D}_{cb} = \mathfrak{C}_e \mathfrak{D}_{21} \mathfrak{B}_t^{-1} \tag{34C}$$
$$\mathfrak{D}_{cc} = \mathfrak{C}_e \mathfrak{D}_{22} \mathfrak{C}_t^{-1} \tag{34D}$$

Thus, it is not necessary to apply the same transformation of variable to the currents and voltages. From equations 34 it appears possible to determine a set of transformations which will result in the preassigned matrices $\mathfrak{D}_{bb}$, $\mathfrak{D}_{bc}$, $\mathfrak{D}_{cb}$, and $\mathfrak{D}_{cc}$. Unfortunately, these four equations can not be solved directly to determine the transformation matrices. However, upon knowing the desired form of the transformed matrix, it is possible to arrive at a useful transformation by trial and error after a thorough study is made of the symmetries of the matrices.

As would be expected, the same transformation can be used on the primary circuits of the induction motor as was used on the synchronous machine, namely

$$\mathfrak{B}_t = \mathfrak{B}_e = \mathfrak{B} \tag{35}$$

where the elements of $\mathfrak{B}$ are given in equation 23.

Since the transformation is to diagonalize the $\mathfrak{D}_{22}$ matrix, the following relation must hold

$$\mathfrak{B}_{22} \mathfrak{C}^{-1} = \mathfrak{C}^{-1} \mathcal{K} \tag{36}$$

where $\mathcal{K}$ is a diagonal matrix. Apparently, there is more than one transformation that will accomplish this purpose. One such transformation is

$$\mathfrak{C}^{-1} = \begin{bmatrix} 1 & \cos\beta \quad \cos 2\beta \ldots\ldots & \cos(n-2)\beta & 1/n \\ \cos(-\Delta) & \cos(\beta-\Delta)\cos(2\beta-\Delta)\ldots & \cos[(n-2)\beta-\Delta] & 1/n \\ \vdots & & \vdots & \\ \cos[-n-1)\Delta] & \ldots\ldots\ldots\ldots\ldots\ldots & \cos[(n-2)\beta-(n-1)\Delta] & 1/n \end{bmatrix} \tag{37}$$

where $\mathfrak{B} = \pi/n-1$ and the $1/n$ terms are introduced into the last column so that $\mathfrak{C}^{-1}$ is nonsingular. Any other constant would serve equally well.

Unfortunately, it has not been possible to evaluate the general form of $\mathfrak{C}$ by taking the inverse of $\mathfrak{C}^{-1}$. Hence, it is necessary to evaluate the inverse for each particular problem. Admittedly, the task of finding the inverse of $\mathfrak{C}^{-1}$ for a squirrel cage with 20 bars may not be practical. Clearly, this particular problem needs further consideration.

For a 3-phase secondary, the transformation matrix given in equation 37 reduces to

$$\mathfrak{C}^{-1} = \begin{bmatrix} 1 & 0 & 1/2 \\ -1/2 & \sqrt{3}/2 & 1/2 \\ -1/2 & -\sqrt{3}/2 & 1/2 \end{bmatrix}$$

and

$$\mathfrak{C} = \frac{2}{3} \begin{bmatrix} 1 & -1/2 & -1/2 \\ 0 & \sqrt{3}/2 & -\sqrt{3}/2 \\ 1 & 1 & 1 \end{bmatrix} \tag{38}$$

By forming the matrix products indicated by this transformation, the detailed expression for matrix equation 34 is

$$\begin{bmatrix} V_d \\ V_q \\ V_o \\ \hline V_a \\ V_b \\ V_c \end{bmatrix} = \left[\begin{array}{ccc|ccc} r_p + L_d P & L_q \dot{\phi} & 0 & \frac{3}{2} M_{rs} P & \frac{3}{2} M_{rs} \dot{\phi} & 0 \\ -L_d \dot{\phi} & r_p + L_q P & 0 & \frac{3}{2} M_{rs} \dot{\phi} & -\frac{3}{2} M_{rs} P & 0 \\ 0 & 0 & r_p + L_o P & 0 & 0 & 0 \\ \hline \frac{3}{2} M_{rs} P & 0 & 0 & r_s + L_a P & 0 & 0 \\ 0 & \frac{3}{2} M_{rs} P & 0 & 0 & r_s + L_b P & 0 \\ 0 & 0 & 0 & 0 & 0 & r_s + L_c P \end{array}\right] \begin{bmatrix} i_d \\ i_q \\ i_o \\ \hline i_a \\ i_b \\ i_c \end{bmatrix}$$

where

r_p = resistance of the primary per phase
r_s = resistance of the secondary per phase
$L_a = L_b = L_s + M_s$
$L_c = L_s - 2M_s$
$L_d = L_q = L_o + M_o$
$L_o = (L_o - M_o)/2$
L_o and M_o represent the self- and mutual inductances between the primary phases.
L_s and M_s represent the self- and mutual inductances between secondary phases.

The transformations used here on the primary and secondary mesh circuits has resulted in a set of constant coefficient differential equations which can be solved both for the transient and steady-state conditions of operation.

Other transformations could have been applied to the original set of differential equations. For example, if only the steady-state solution of the equations is desired, the same transformation of variable could have been used on the primary circuits as was used on the seconday circuits. In this case the transformation leads easily and quickly to the familiar cross-field theory of polyphase induction machines. The cross-field theory of single-phase machines is obtained by applying the $\mathfrak{C}$ matrix transformation to the secondary of the induction machine, with no transformation applied to the primary circuit.

If the voltages applied to the primary of a 3-phase induction machine are unbalanced, the symmetrical-component transformation is a convenient one to use. However, this transformation can not be used until the equations have been reduced to linear functions in sin *ωt*.

Summary

From the point of view presented in this paper, the problem of analyzing rotating machinery appears as a network problem in which there is relative motion between some or all of the elements of the machine. With only slight modification, standard network theory can be applied to the analysis of this type of network. The general procedure to be used in the analysis of rotating machinery is therefore very similar to, and can be outlined as precisely as the procedure used in the analysis of static networks. Thus, it is possible to teach a method of machine analysis, or perhaps more correctly, a method of dynamic network analysis, just as a method of static network analysis is currently taught. A brief outline of the steps to be taken in the network analysis of rotating machinery can be stated as follows:

1. Establish a map or diagram showing the connections between the elements of the machine.

2. Establish a map or reference frame from which the inductances of the elements of the machine can be determined as a function of the angular displacement between the elements.

3. Using standard network theory, establish the differential equations of the network being careful to take the derivative of all inductance-current products.

4. Where possible, apply a transformation of variable to remove as many variable coefficients as possible or diagonalize the matrix containing the coefficients of the differential equations. If only the steady-state solution is required, the variable-coefficient differential equations can usually be solved without applying a transformation of variable.

5. Evaluate the coefficients of the equations and obtain the solution. With the proper interpretation, existing methods can be used to evaluate the coefficients.

It has been shown that most of the concepts used in the classical methods of machine analysis can be obtained from this point of view with little or no added difficulty. Thus, by expanding present concepts of circuit analysis and taking a slightly different point of view toward the rotating-machinery problem, these two areas of technology can be built on a common and rigorous foundation.

References

1. Application of Network Theory to the Analysis of Rotating Machinery. Part II—Commutating Machines, H. E. Koenig. *AIEE Transactions*, vol. 73, pt. I, May 1954, pp. 169–77.

2. The Application of Tensors to the Analysis of Rotating Electrical Machinery, **Gabriel Kron.** *General Electric Review*, Schenectady, N. Y., pts. I–XVI, 1938.

3. The Modern Approach to Electrical Machine Analysis, **W. J. Gibbs.** *The Engineer*, London, England, Oct. 12, 19, 26, and Nov. 2, 1951.

4. Matrix Analysis of Unsymmetrical Machines, **S. Ganapathy, C. S. Ghosh.** *Journal of Indian Institute of Science*, Bangalore, India, vol. XXXV, No. 2, 1953.

5. Fundamental Concepts of Synchronous Machine Reactances, **B. R. Prentice.** *AIEE Transactions (Electrical Engineering)*, vol. 56, 1937, Supplement, pp. 1–21.

(See discussion on page 174)

Application of Network Theory to the Analysis of Rotating Machinery

Part II—Commutating Machines

H. E. KOENIG
ASSOCIATE MEMBER AIEE

THE material presented in this paper is intended to be a continuation of the rotating machinery theory presented in reference 1. The reader will find it desirable to refer to this and the conference paper, "General Network Theory in Terms of Matrix Algebra," presented by M. B. Reed at the 1954 AIEE Winter General Meeting.

When network theory is used, the analysis of commutating machines differs from that of the noncommutating machines in that a new transient may be initiated every time a commutating segment enters or leaves the brush contact area. Thus, to obtain the complete solution it is necessary to solve the problem for each of the intervals of time during which no switching occurs until the solutions repeat. The final value of the current in any given interval becomes the initial values of the currents in the succeeding interval.

The material presented in this paper is not intended to be an exhaustive treatment of the commutating machine problem. The purpose is rather to outline a method whereby network theory can be used to study the commutation characteristics as well as the external characteristics of the machine and to outline a method of processing the differential equations of the machine to reduce their complexity to a point where they have practical engineering value. When the coefficients in the resulting differential equations are evaluated and the equations solved with the aid of computing equipment, further understanding of the commutation phenomenon and the conditions necessary to obtain good commutation both under transient and steady-state conditions will be obtained.

To aid the reader in understanding the general procedure and the point of view presented, a single-commutator d-c machine with a simplex lap-wound armature is considered as an example. The more complex problems can be treated in the same manner, but owing to the added number of circuits, additional complexity is sure to result. It is perhaps safe to say that a problem as complex as that of the commutating machine could not be comprehended without the aid of matrix algebra. Among other things, the use of this tool makes it possible to subdivide a very complex set of linear equations into small subsets that can be readily comprehended and analyzed.

Assumptions, Reference Systems, and Conventions

When any problem as complex as that of the d-c machine is to be analyzed, certain assumptions are necessary. In this treatise the machine is considered to possess the following properties:

1. All inductances are constant over the operating range considered;
2. All portions of the machine except the frame are to be free of eddy currents;
3. Capacitance between conductors is neglected.

As in the analysis of the noncommutating machines, the analysis of the commutating machine is considered in three major parts, namely:

1. Establish a set of differential equations describing the performance of the machine in terms of circuit parameters.
2. Evaluate the circuit parameters in terms of machine dimensions.
3. Solve the resulting equations to obtain the general transient or frequency response.

Only the first part of the problem is considered in this paper.

The circuit conventions and reference systems used are consistent with those used in reference 1.

Nomenclature

Nomenclature not defined in the paper or which is not obvious from the network diagrams is defined as follows:

a_1 = number of elements in armature branch circuit 1, an element being defined as that portion of the armature winding between commutator segments
a_2 = number of elements in armature branch circuit 2
n = total number of armature elements
q_1 = number of elements short circuited by brush 1
q_2 = number of elements short circuited by brush 2
$n_1 = q_1 + a_1$
$P = d/dt$
ϕ = angular position of rotor
e_a = terminal voltage of armature
e_f = terminal voltage of shunt field
r_j, L_j = resistance and inductance of jth armature element
M_{jk} = magnetic coupling between jth and kth armature element

---

Paper 54-88, recommended by the AIEE Basic Sciences and Rotating Machinery Committees and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 19, 1953; made available for printing December 4, 1953.

H. E. Koenig is with the University of Illinois, Urbana, Ill.

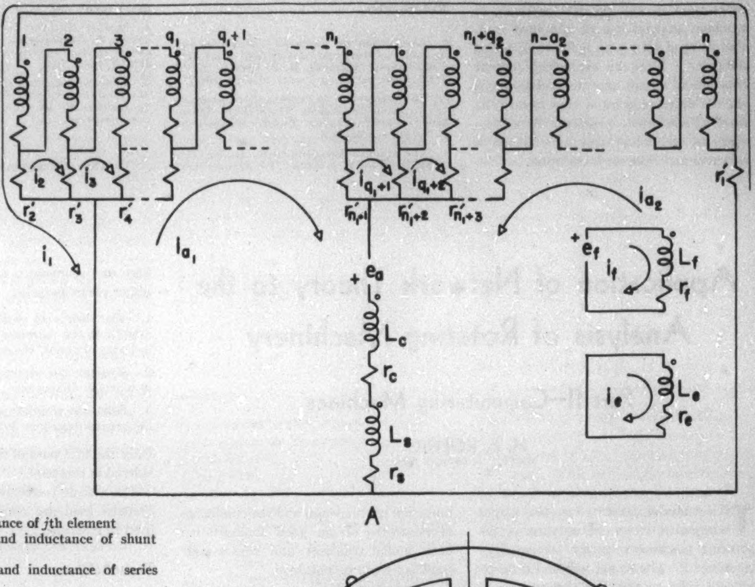
Acknowledgment is made to the Graduate College of the University of Illinois for permission to utilize material in a theses, "A Circuit Approach to the Analysis of Electrical Machinery," presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical Engineering.

The author also acknowledges the encouragement and material assistance received from the Reliance Electric and Engineering Company and Doctors G. H. Fett and M. B. Reed, Electrical Engineering Department, University of Illinois.

Fig. 1. D-c machine

A. Connection diagram

B. Map showing physical orientation of machine elements



r_j' = contact resistance of jth element

r_f, L_f = resistance and inductance of shunt field

r_s, L_s = resistance and inductance of series field

r_c, L_c = resistance and inductance of commutating pole field

r_e, L_e = equivalent resistance and inductance of eddy currents in frame

Differential Equations of Commutating Machines

If the commutating phenomenon of the machine is to be included in the analysis, it is necessary to consider each interval of time during which a given number of commutating segments are short-circuited by the brushes. For each of these intervals, a connection diagram for the machine can be established from which the differential equations can be written and solved. By considering the general case of an arbitrary number of coils short-circuited by each brush, a set of equations are derived which apply to each of the intervals to be considered. The general procedure for establishing the necessary differential equations and some of the techniques of manipulation are next demonstrated for a machine having a simplex lap-wound armature with an integral number of commutator segments per pair of poles. The same procedure is applicable to other types of winding, but more complexity may result. Since the winding to be considered in this example repeats every pair of poles, the analysis can be made on the basis of one pair of poles.

When the portions of the armature winding between successive commutator segments are considered as elements of the machine, the connection diagram for a machine with n commutator segments per pair of poles becomes that shown in Fig. 1(A), where all symbols used on the diagram are defined in the Nomenclature. For example the contact resistances are represented by $r_1', r_2', \ldots, r_j'$. Fig. 1(B) is a map or diagram from which the inductances appearing in Fig. 1(A) are determined as a function of rotor position. The effect of the eddy currents in the solid iron frame are approximated by a closed winding e. For sinusoidally time-varying currents, the equivalent inductance and resistance of this winding can be evaluated by solving Maxwell's field equations in the frame region. If the frame is considered as a semi-infinite iron sheet, the solution is not difficult to obtain and is very similar to the solution of the wave guide problems. However, for other than sinusoidally time-varying currents the solution to Maxwell's equations in the frame region is very difficult, if not impossible, to obtain in closed form. Thus it may be necessary to study the transient commutating ability of solid-frame machines in terms of frequency characteristics if any degree of accuracy is to be obtained. For the type of material used in the usual d-c machine, the depth of penetration is of the order of 1 centimeter when the frequency is only 1 cycle per second.

If the brushes span several segments and if it is necessary to include more than one pair of poles in the analysis, the connection diagram of the machine becomes very complex, as does writing the mesh-circuit differential equations from inspection of the diagram. In this case the matrix transformation of variable should be used to transform the branch circuit equations into a set of mesh-circuit variables. This technique is outlined in M. B. Reed's paper on the general subject of network theory. By following this procedure or from inspection of the connection diagram of Fig. 1(A), the differential equations for the d-c machine can now be written. In matrix notation, the set of equations for Fig. 1(A) are represented by

$$\mathcal{E} = \mathfrak{D}\mathcal{I} \qquad (1)$$

where the $\mathcal{E}$ matrix contains the sum of the driving voltages around the successive mesh circuits, $\mathfrak{D}$ contains the coefficients in the differential equations, and $\mathcal{I}$ contains the unknown mesh-circuit currents. Before attempting to show explicitly what the elements of these matrices are, it is convenient to partition the matrices of equation 1 so that the submatrices can be manipulated with ease. In general, the partitioning is quite arbitrary so long as the partitioning is conformal. However, as becomes apparent, it is convenient to partition the matrices in a special manner. Let matrix equation 1 be written

$$\begin{bmatrix} \mathcal{E}_1 \\ \mathcal{E}_2 \\ \mathcal{E}_3 \\ \mathcal{E}_4 \end{bmatrix} = \begin{bmatrix} \mathfrak{D}_{11} & \mathfrak{D}_{12} & \mathfrak{D}_{13} & \mathfrak{D}_{14} \\ \mathfrak{D}_{21} & \mathfrak{D}_{22} & \mathfrak{D}_{23} & \mathfrak{D}_{24} \\ \mathfrak{D}_{31} & \mathfrak{D}_{32} & \mathfrak{D}_{33} & \mathfrak{D}_{34} \\ \mathfrak{D}_{41} & \mathfrak{D}_{42} & \mathfrak{D}_{43} & \mathfrak{D}_{44} \end{bmatrix} \begin{bmatrix} \mathcal{I}_1 \\ \mathcal{I}_2 \\ \mathcal{I}_3 \\ \mathcal{I}^4 \end{bmatrix} \qquad (2)$$

The set of terms appearing in each of these submatrices are shown explicitly in the Appendix. The experienced person can fill in the coefficients of these matrices from inspection of the diagram. The uninitiated may find it desirable or necessary to first write the differential equations for the network in the usual manner and then arrange the coefficients in matrix form. Briefly, the voltage matrices $\mathcal{E}_1$ and $\mathcal{E}_2$ represent the driving voltages in the mesh circuits short-circuited by brushes 1 and 2, respectively. The voltages of $\mathcal{E}_3$ represent the driving voltages in the two main branch circuits of the armature. The $\mathcal{E}_4$ matrix represents the driving voltages in the field circuits.

The coefficients of the $\mathfrak{D}$ matrices, as shown in the Appendix, are all expressed in terms of the inductance and resistance of each element of the machine. Thus, the inductance of the main branches of the armature, L_{a_1}, and L_{a_2}, for example, are expressed in terms of the self and mutual inductances of each element of the armature winding. Once the coupling between the jth and kth elements of the armature is determined as a general function of rotor position, the summations indicated in the Appendix can be readily evaluated to obtain all the inductance coefficients of the armature as a function of rotor position.

The resulting set of variable-coefficient differential equations represented by matrix equation 2 are very difficult, if not impossible, to solve in this particular form. If all harmonics in the inductances above the second could be neglected, it would be possible to effect a transformation of variable to remove at least part of the variable coefficients, as was done in the analysis of the noncommutating machines discussed in M. B. Reed's paper. If harmonics above the second can not be neglected, and in the usual d-c machine they can not, the desired transformation may be very difficult to obtain.

The other approach to the solution of the equations is a point-by-point solution. Under the original specifications, the equations obtained thus far in this discussion only apply during a short interval of time, during which a given number of armature elements are short-circuited by the brushes. Thus, as a first approximation, the variable coefficients can be considered constant for the short intervals of time being considered. If a higher degree of accuracy is required, shorter intervals of time can be considered, or, as is shown subsequently, the solution can be considered in parts and the complexity of the equations reduced to the point where the variable coefficients can be included in the solution.

In analyzing commutating machinery there are roughly two major problems to be considered:

1. Assuming the machine will commutate satisfactorily, what are the external transient and steady-state characteristics of the machine when used as a component of an electrical system and subjected to a given set of operating conditions?
2. Knowing the external characteristics and the conditions of operation, will the machine commutate satisfactorily, and if not, what must be done to insure satisfactory commutation?

From this point of view one is not particularly interested in analyzing the machine under very poor conditions of commutation, where severe arcing conditions make it impossible to evaluate the contact resistances. The techniques used in processing the differential equations of the machine to obtain the answers to these two questions will be considered in the order given.

External Characteristics

To reduce the complexity of these equations to a point where they have practical application, it is necessary to make several approximations. In the $\mathfrak{D}_{31}$ and $\mathfrak{D}_{32}$ matrices, the terms representing the magnetic coupling between the armature branch circuits and the jth coil in the commutating zone, M_{j,a_1} and M_{j,a_2}, are negligibly small, since the axes of the coils are at right angles. As in most analyses of d-c machines, these terms, as well as the resistance terms appearing in these two matrices are neglected. The $M_{j,a_3} = M_{j,c} + M_{j,s}$ terms in these same two matrices represent the magnetic coupling between the jth coil in the commutating zone and the two field circuits connected in series with the main circuit of the armature (commutating field and compounding fields). Since these inductances are functions of time, the derivative of the current-inductance product contains two terms of the form

$$P(M_{j,f}i_j) = M_{j,f}Pi_j + i_jPM_{j,f} \qquad (3)$$

where the two resulting terms represent, respectively, the transformer and rotational components of voltage induced in the field circuits by the current in the jth coil in the commutating zone. As symmetric commutation is approached, the total of all the components of rotational voltages in the field windings, as produced by the currents in the coils in the commutating zone, approaches zero. The $\mathfrak{D}_{41}$ and $\mathfrak{D}_{42}$ matrices contain similar terms to which the same argument applies. Thus, all terms of the form $i_j(PM_{j,f})$ appearing in the $\mathfrak{D}_{31}$, $\mathfrak{D}_{32}$, $\mathfrak{D}_{41}$, and $\mathfrak{D}_{42}$ matrices can be neglected. The only terms remaining in these four matrices are those of the form, $M_{j,f}P$, which when multiplied by the current in the coils being commutated i_j give terms of the form $M_{j,f}Pi_j$. These terms represent components of voltage in the field windings resulting from the changing currents in the commutating zone. For the main field windings, these components of voltage are indeed large, since the axes of the coils are nearly parallel and the coefficient of coupling high. Preliminary calculations on a 200-horsepower 230-volt 1,150-rpm d-c motor indicate that for linear commutation, the total voltage induced in the shunt field by the currents in the coils undergoing commutation is in the order of 20,000 volts.

However, these effects are counterbalanced, for the most part, by the effects of armature reaction. In the $\mathfrak{D}_{33}$ and $\mathfrak{D}_{43}$ there are terms of the form $(PM_{a,f})$ which, when multiplied by the currents i_a in the main branch circuits of the machine, give terms of the form $(PM_{a,f})$ i_a. These terms represent components of voltages induced in the field windings by the rotation of the armature reaction field during the interval of time during which a given number of coils are short circuited by the brushes. It can be shown that in the windings of the main poles, the sum of the two components of voltages discussed in the foregoing vanishes as linear commutation is approached. For anything but linear commutation, there are resulting high-frequency components of voltages induced in the windings of the machine as it commutates. Perhaps these voltages are a source of some of the radio-noise experience in d-c machines.

The information concerning the secondary effects discussed in the previous paragraph are removed from the differential equations of the machine by neglecting the coefficients in the $\mathfrak{D}_{31}$, $\mathfrak{D}_{32}$, $\mathfrak{D}_{41}$, and $\mathfrak{D}_{42}$ matrices, setting the (PM_{a_1f}), (PM_{a_1e}), (PM_{a_2e}), and (PM_{a_2f}) terms in the $\mathfrak{D}_{43}$ matrix equal to zero, and dropping one of the $(PM_{a_1a_0})$ or $(PM_{a_2a_0})$ terms from each of the four elements of the $\mathfrak{D}_{33}$ matrix. Note that no terms are to be removed from the $\mathfrak{D}_{34}$ matrix. The currents in the main branch circuits of the armature and the field windings are now given by the solution of the set of differential equations represented by the following matrix equation

$$\begin{bmatrix} \mathcal{E}_3 \\ \mathcal{E}_4 \end{bmatrix} = \begin{bmatrix} \mathfrak{D}_{33}^* & \mathfrak{D}_{34} \\ \mathfrak{D}_{43}^* & \mathfrak{D}_{44} \end{bmatrix} \begin{bmatrix} \mathcal{I}_3 \\ \mathcal{I}_4 \end{bmatrix} \tag{4}$$

where the $\mathfrak{D}_{33}^*$ and $\mathfrak{D}_{43}^*$ matrices are the modified $\mathfrak{D}_{33}$ and $\mathfrak{D}_{43}$ matrices.

Since the set of differential equations represented by matrix equation 4 applies only for a short interval of time, further simplification results if the coefficients of the differential equations are assumed constant after the derivative of the current-inductance products are expanded. It is also permissible to assume the two branches of the armature circuit to be identical, and assume the armature branch circuits equal. With these additional simplifications, the differential equations reduce to

$$\begin{bmatrix} e_a \\ e_f \\ 0 \end{bmatrix} = \begin{bmatrix} (R_a+\phi M_{a_1a_0}) & \phi M_{a_1f} & \phi M_{a_1e} \\ 0 & r_f & 0 \\ 0 & 0 & r_e \end{bmatrix} \begin{bmatrix} i_1 \\ i_f \\ i_e \end{bmatrix} + \begin{bmatrix} 2(L_{a_1}+2M_{a_1a_0}) & (M_{a_1f}+M_{sf}) & (M_{a_1e}+M_{se}+M_{ce}) \\ (M_{a_1f}+M_{sf}) & L_f & M_{fe} \\ (M_{a_1e}+M_{se}+M_{ce} & M_{fe} & L_e \end{bmatrix} P \begin{bmatrix} {}_1i \\ i_f \\ i_e \end{bmatrix} \tag{5}$$

where $R_a=(r_{a_1+1}+r_{a_1}+r_{n_1+1}+2r_{a_0})$ and all other terms have been previously defined.

If it is further assumed that the number of coils short-circuited by each of the brushes is the same for all intervals of time, then the network is the same for all intervals of time, and the solution to these equations represents the approximate solution to the currents in the armature and fields as a continuous function of time.

Admittedly, similar equations for the d-c machine have been arrived at by other authors using methods that may appear less complex to those unfamiliar with matrix notation. However, it is believed that the additional information concerning the approximations necessary to arrive at this end result justifies this apparent added complexity, to say nothing of the commutation study which results from this approach to the problem and which is considered next.

Commutating Characteristics

The complexity of the equations of commutation are greatly reduced if the armature and field currents, $\mathcal{I}_3$ and $\mathcal{I}_4$, are first obtained by solving the set of differential equations represented by matrix equation 5. These currents are then substituted into matrix equation 2 as known functions. Observing that $\mathcal{E}_1 \equiv \mathcal{E}_2 \equiv 0$, the set of differential equations describing the commutating characteristics reduce simply to

$$-\begin{bmatrix} \mathfrak{D}_{13} & \mathfrak{D}_{14} \\ \mathfrak{D}_{23} & \mathfrak{D}_{24} \end{bmatrix} \begin{bmatrix} \mathcal{I}_3 \\ \mathcal{I}_4 \end{bmatrix} = \begin{bmatrix} \mathfrak{D}_{11} & \mathfrak{D}_{12} \\ \mathfrak{D}_{21} & \mathfrak{D}_{22} \end{bmatrix} \begin{bmatrix} \mathcal{I}_1 \\ \mathcal{I}_2 \end{bmatrix} \tag{6}$$

where the known current functions, $\mathcal{I}_3$ and $\mathcal{I}_4$, appear as driving voltages in equations of the coils in the commutating zone. The number of differential equations represented by matrix equation 6 is of course equal to the number of coils short-circuited by the two brushes. The elements of the $\mathfrak{D}$ matrices to the right of the equality sign represent the resistances and inductances of the coils in the commutating zone, and the coupling between them. These coefficients are all of the same order of magnitude, are nonlinear, and are functions of rotor position. As a general proposition, considerable improvement in accuracy is obtained in the study of such a set of equations if a transformation of variable can be effected which results in a set of equations having the differences between the self- and mutual inductances as coefficients. It is shown[1] that the 2-winding transformer problem can be treated in this manner, and the usual **T**-equivalent obtained by applying a transformation of variable to the original coupled circuit equations of the transformer. Clearly, the objective in these two problems is identical. There are several transformations that will accomplish the desired objective. The general procedure for finding this type of transformation is shown next.

Let a new set of driving voltages and currents be defined as:

$$\begin{aligned} &\mathcal{I}_a=\mathfrak{C}_1^{-1}\mathcal{I}_1;\ \ \mathcal{I}_b=\mathfrak{C}_2^{-1}\mathcal{I}_2 \\ &\mathcal{E}_a=\mathfrak{C}_1'\mathcal{E}_1^*;\ \ \mathcal{E}_b=\mathfrak{C}_2'\mathcal{E}_2^* \end{aligned} \tag{7}$$

where the $\mathfrak{C}_1$ and $\mathfrak{C}_2$ matrices are yet to be determined and the driving voltages to the left of the equality sign in equation 6 are represented simply by $\mathcal{E}_1^*$ and $\mathcal{E}_2^*$. Upon applying this transformation of variable to equation 6 there results

$$\begin{bmatrix} \mathcal{E}_a \\ \mathcal{E}_b \end{bmatrix} = \begin{bmatrix} \mathfrak{C}_1'\mathfrak{D}_{11}\mathfrak{C}_1 & \mathfrak{C}_1'\mathfrak{D}_{12}\mathfrak{C}_2 \\ \mathfrak{C}_2'\mathfrak{D}_{21}\mathfrak{C}_1 & \mathfrak{C}_2'\mathfrak{D}_{22}\mathfrak{C}_2 \end{bmatrix} \begin{bmatrix} \mathcal{I}_a \\ \mathcal{I}_b \end{bmatrix} \tag{8}$$

With a little practice and by observing that the elements in each row of the $\mathfrak{D}_{11}$ matrix, for example, are multiplied by the corresponding elements in the corresponding columns of the $\mathfrak{C}_1$ matrix, it is found that 1's, −1's or 0's may be inserted in the rows of the $\mathfrak{C}_1$ matrix in such a manner as to add or subtract any pair or pairs of columns in the $\mathfrak{D}_{11}$ matrix when the product is formed. When the product $\mathfrak{C}_1'(\mathfrak{D}_{11}\mathfrak{C}_1)$ is formed, the corresponding rows in the $\mathfrak{D}_{11}\mathfrak{C}_1$ matix are added or subtracted. It is easy to show that this type of transformation always results in a symmetric matrix if the $\mathfrak{D}_{11}$ matrix is symmetric.

In general, the elements of the transformation matrices depend entirely upon the particular problem being considered, and can be any combination of numbers or functions as long as the transformation matrix has an inverse. To show a specific example, suppose there are three coils short circuited by each brush of the machine, and that the first two coils under the first brush occupy the same slot. The elements in the $\mathfrak{D}_{11}$ matrix are given by equation 11 of the Appendix, with $M_{11}=M_{22}=M_{12}$ and $M_{13}=M_{23}$, if perfect coupling is assumed. In this case, a useful transformation is

$$\mathfrak{C}=\begin{bmatrix}1 & -1 & -1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} \tag{9}$$

By forming the indicated matrix products the inductance coefficients in the matrix take on the form

$$P\begin{bmatrix}M_{11} & 0 & (M_{13}-M_{11})\\ 0 & 0 & 0\\ (M_{31}-M_{11}) & 0 & (M_{11}+M_{33}-2M_{13})\end{bmatrix}$$

where much improvement in accuracy is obtained and the coefficients are more nearly linear. The differences in the inductances of the coils, normally referred to as the leakage inductances, can now be evaluated instead of the inductance of each coil. It is also clear that the complexity of the equations is greatly reduced as a result of the zeros in the transformed matrix.

After processing the differential equations of the machine in a manner similar to that outlined above, the differential equations describing the commutating characteristics reduce to a practical set of equations, the solution of which can be obtained with the aid of computing equipment. When these equations are solved for each interval of commutation, the currents in the coils during that interval are obtained, the final value of the current in any particular interval serving as initial conditions for the succeeding interval. Clearly, certain assumptions must be made concerning the contact resistances. However, it is believed that the contact resistance voltages are small compared to the inductance voltages, so inaccuracies caused by the uncertainties in the value of this term are small.

Summary

The procedure used in arriving at the differential equations of the commutating machine is identical to that used on the noncommutating machines discussed in reference 1. However, the techniques used in processing the resulting equations is somewhat different. In the study of commutating machines, two problems are of interest:

1. How can the machine be analyzed along with its associated network to determine the external characteristics?
2. Will the machine commutate satisfactorily under the conditions of operation?

The answer to the first question is obtained by neglecting, in as much as is practical, the effect of the currents in the coils undergoing commutation when solving for the external characteristics. In so doing, the complexity of the problem is reduced to the point where it is practical to solve the equations of the machine and its associated network to determine the armature and field currents. The known armature and field currents are then substituted into the differential equations as driving functions and a study made of the commutating characteristics. Much improvement in accuracy and considerable simplification is obtained by applying a transformation of variable, which produces a set of differential equations whose coefficients are the difference between the self- and mutual inductances of the armature coils. It is perhaps safe to say that without the use of matrix algebra, a network problem as complex as that of the commutating machine would be very difficult, if not impossible, to comprehend.

Appendix

Matrix Elements

The elements of each of the submatrices in matrix equation 2 can be expressed in terms of the resistances and inductances of each of the elements of the machine. It is standard practice to write the mesh-circuit equations in the sequence of the increasing subscripts on the mesh currents. When this procedure is followed, the elements of the submatrices of equation 2 as obtained from the differential equations of the network are

$$\mathcal{E}_1\equiv 0,\ \mathcal{E}_2\equiv 0,\ \mathcal{E}_3=\begin{bmatrix}e_a\\ e_a\end{bmatrix}\quad \mathcal{E}_4=\begin{bmatrix}e_f\\ 0\end{bmatrix} \tag{10}$$

$$\mathfrak{D}_{13}=\mathfrak{D}_{31}'=\begin{bmatrix}0 & r_1'\\ 0 & 0\\ \cdot & \cdot\\ \cdot & \cdot\\ \cdot & \cdot\\ -r_{q_1+1}' & 0\end{bmatrix}+$$

$$P\begin{bmatrix}(M_{1,a_1}+M_{1,a_0}) & (-M_{1,a_2}+M_{1,a_0})\\ (M_{2,a_1}+M_{2,a_0}) & (-M_{2,a_2}+M_{2,a_0})\\ \cdot & \cdot\\ \cdot & \cdot\\ \cdot & \cdot\\ (M_{q_1,a_1}+M_{q_1}+{}_{a_0}) & (-M_{q_1,a_2}+M_{q_1,a_0})\end{bmatrix} \tag{13}$$

where

$$M_{k,a_1}=\sum_{j=q_1+1}^{q_1+a_1} M_{k,j}$$

$$M_{k,a_2}=\sum_{j=n-a_2}^{n} M_{k,j}$$

$$M_{k,a_0}=M_{k,c}+M_{k,s}$$

$$\mathfrak{D}_{14}=\mathfrak{D}_{41}'=P\begin{bmatrix}M_{1,f} & M_{1,e}\\ M_{2,f} & M_{2,e}\\ \cdot & \cdot\\ \cdot & \cdot\\ \cdot & \cdot\\ M_{q_1,f} & M_{q_1,e}\end{bmatrix} \tag{14}$$

$$\mathfrak{D}_{22}=$$

$$\begin{bmatrix}(r_{n_1+1}'+r_{n_1+2}'+r_{n_1+1}') & -r_{n_1+2}' & \cdots\\ -r_{n_1+2}' & (r_{n_1+2}'+r_{n_1+3}'+r_{n_1+2}) & \cdots\\ \cdot & \cdot & \\ \cdot & \cdot & \\ \cdot & \cdot & \end{bmatrix}+$$

$$P\begin{bmatrix}M_{n_1+1,n_1+1} & M_{n_1+1,n_1+2}\cdots\\ M_{n_1+2,n_1+1} & M_{n_1+2,n_1+2}\cdots\\ \cdot & \cdot\\ \cdot & \cdot\\ \cdot & \cdot\end{bmatrix} \tag{15}$$

$$\mathfrak{D}_{11}=\begin{bmatrix}(r_1'+r_2'+r_1) & -r_2' & 0 & \cdots\\ -r_2' & (r_2'+r_3'+r_2) & -r_3' & \cdots\\ 0 & -r_3' & (r_3'+r_4'+r_3) & \cdots\\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \end{bmatrix}+P\begin{bmatrix}M_{11} & M_{12} & M_{13} & \cdots\\ M_{21} & M_{22} & M_{23} & \cdots\\ M_{31} & M_{32} & M_{33} & \cdots\\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \end{bmatrix} \tag{11}$$

$$\mathfrak{D}_{12}=\mathfrak{D}_{21}'=-P\begin{bmatrix}M_{1,n_1+1} & M_{1,n_1+2} & M_{1,n_1+3}\cdots\\ M_{2,n_1+1} & M_{2,n_1+2} & M_{2,n_1+3}\cdots\\ M_{3,n_1+1} & M_{3,n_1+2} & M_{3,n_1+3}\cdots\\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\\ \cdot & \cdot & \cdot\end{bmatrix} \tag{12}$$

$$\mathfrak{D}_{22}=\mathfrak{D}_{22}'=\begin{bmatrix} r_{n_1+1}' & 0\\ 0 & 0\\ \cdot & \cdot\\ \cdot & \cdot\\ \cdot & \cdot \end{bmatrix}-P\begin{bmatrix}(M_{n_1+1,a_1}+M_{n_1+1,a_0})(-M_{n_1+1,a_2}+M_{n_1+1,a_0})\\ (M_{n_1+2,a_1}+M_{n_1+2,a_0})(-M_{n_1+2,a_2}+M_{n_1+2,a_0})\\ \cdot \qquad\qquad \cdot\\ \cdot \qquad\qquad \cdot\\ \cdot \qquad\qquad \cdot\end{bmatrix} \tag{16}$$

where

$n_1=q_1+a_1$

$$M_{k,a_1}=\sum_{j=q_1+1}^{q_1+a_1} M_{k,j}$$

$$M_{k,a_2}=\sum_{j=n-a_2}^{n} M_{k,j}$$

$$M_{k,a_0}=M_{k,c}+M_{k,s}$$

$$\mathfrak{D}_{24}=\mathfrak{D}_{42}'=-P\begin{bmatrix} M_{n_1+1,f} & M_{n_1+1,e}\\ M_{n_1+2,f} & M_{n_1+2,e}\\ \cdot & \cdot\\ \cdot & \cdot\\ \cdot & \cdot \end{bmatrix} \tag{17}$$

$$M_{a_1a_2}=\sum_{j=q_1+1}^{q_1+a_1}\ \sum_{k=n-a_2}^{n} M_{j,k}$$

$$L_{a_0}=L_c+L_s+2M_{cs}$$

$$M_{a_1a_0}=\sum_{j=q_1+1}^{q_1+a_1}(M_{j,s}+M_{j,c})$$

$$M_{a_2a_0}=\sum_{j=n-a_2}^{n}(M_{j,s}+M_{j,c})$$

$$\mathfrak{D}_{33}=\begin{bmatrix}(r_{q_1+1}'+r_{a_1}+r_{n_1+1}'+r_{a_0})+r_{a_0} & +r_{a_0}\\ +r_{a_0} & (r_{n-a_2}'+r_{a_2}+r_1'+r_{a_0})\end{bmatrix}+$$

$$P\begin{bmatrix}(L_{a_1}+2M_{a_1a_0}+L_{a_0}) & (L_{a_0}-M_{a_2a_0}+M_{a_1a_0}-M_{a_1a_2})\\ (L_{a_0}-M_{a_2a_0}+M_{a_1a_0}-M_{a_1a_2}) & (L_{a_2}+2M_{a_2a_0}+L_{a_0})\end{bmatrix} \tag{18}$$

where

$$L_{a_1}=\sum_{j=q_1+1}^{q_1+a_1}\ \sum_{k=q_1+1}^{q_1+a_1} M_{j,k}$$

$$L_{a_2}=\sum_{j=n-a_2}^{n}\ \sum_{k=n-a_2}^{n} M_{j,k}$$

$$\mathfrak{D}_{34}=\mathfrak{D}_{43}'$$

$$=P\begin{bmatrix}(M_{a_1,f}+M_{a_0,f}) & (M_{a_1,e}+M_{a_0,e})\\ (M_{a_0,f}-M_{a_2,f}) & (M_{a_0,e}-M_{a_2,e})\end{bmatrix} \tag{19}$$

where

$$M_{a_1,f}=\sum_{j=q_1+1}^{q_1+a_1} M_{j,f}$$

$$M_{a_1,e}=\sum_{j=q_1+1}^{q_1+a_1} M_{j,e}$$

$$M_{a_0,f}=M_{c,f}+M_{s,f}$$

$$M_{a_2,f}=\sum_{j=n-a_2}^{n} M_{j,f}$$

$$M_{a_2,e}=\sum_{j=n-a_2}^{n} M_{j,e}$$

$$M_{a0,e}=M_{c,e}+M_{s,e}$$

$$\mathfrak{D}_{44}=\begin{bmatrix} r_f & 0\\ 0 & r_e\end{bmatrix}+P\begin{bmatrix} L_f & M_{f,e}\\ M_{f,e} & L_e\end{bmatrix} \tag{20}$$

$$\mathfrak{I}_1=\begin{bmatrix} i_1\\ i_2\\ \cdot\\ \cdot\\ \cdot\\ i_{q_1}\end{bmatrix};\ \mathfrak{I}_2=\begin{bmatrix} i_{q_1+1}\\ i_{q_1+2}\\ \cdot\\ \cdot\\ \cdot\\ i_{q_1+q_2}\end{bmatrix};$$

$$\mathfrak{I}_3=\begin{bmatrix} i_{a_1}\\ i_{a_2}\end{bmatrix};\ \mathfrak{I}_4=\begin{bmatrix} i_f\\ i_e\end{bmatrix}$$

Reference

1. Application of Network Theory to the Analysis of Rotating Machinery. Part I—Synchronous and Asynchronous Machines, H. E. Koenig. *AIEE Transactions*, vol. 73, pt. I, May 1954, pp. 162-69.

◆

Discussion

Robert M. Saunders (University of California, Berkeley, Calif.): My discussion is divided into two parts: first, some particular comments upon the contents of the paper and second, some philosophical comments on the application of Prof. Koenig's methods to the field of teaching.

Regarding the five idealizations made by the author, it should be emphasized that not all of these are necessary restrictions but only practical ones which have been imposed to make the presentation feasible. Today much of our preoccupation concerning machines has to do with such phenomena as saturation, effects of slots, and contributions of magnetomotive force or flux harmonics to voltages and torques. Thus, if his methods are to be at all useful, some of these contraints must be relaxed at times. It would be well if the author reassured us that such as the case, pointing out which may and which may not be discarded and under what conditions.

The author gives certain criteria for ascertaining the sign of the mutually induced voltage term when analyzing circuits. Magnetic polarization marks are not completely arbitrary as the author states. Convention and the American Standards dictate the exact location of polarity marks on mutual inductances and transformers. While I realize that machines have the added element of rotation, I would plead strongly for as much co-ordination between transformer and rotating-machine polarity designation as is possible. In addition it should be possible to ascertain the polarity of the induced voltage solely by recourse to the terminals of coils without needing to know how the coils are wound. Hence, I feel that the criteria given by the author should be revised or clarified.

Toward clarification I should like to propose the following as a substitute for the last paragraph in the section entitled *Assumptions, Reference Systems, and Definitions* in part I:

> Magnetic polarization marks are used on all coils of the machine, their location with respect to each terminal being fixed by standard tests.[1] In ascertaining the location of the polarity marks for coils which may move with respect to each other, the marks should be designated for the coils located at the zero position. Thus, the sign associated with the mutual inductance will always be positive.* The sign of the $d/dt(M_{jk}i_k)$ term is entered into the voltage in accordance with the direction of the induced voltage given by the polarity marks and the rate of change of the kth current when the coils are located at their reference position.

To add a philosophical note, those of us who are educators should have more than a passing interest in this material. The approach of the paper has real significance for it utilizes the direct application of well-known physical principles and rigorous mathematical concepts. For classroom presentation, it is most important to build upon the student's previous knowledge. Thus Prof. Koenig's methods of machine analysis have the powerful background of physics and mathematics upon which to build; this leads to a most effective presentation. He is to be warmly congratulated by those of us who are concerned with teaching electromechanical energy conversion principles at the graduate level.

REFERENCE

1. Test Code for Distribution, Power, and Regulating Transformers. *Publication No. C57.22*, American Standards Association, New York, N. Y., 1948.

C. G. Veinott (Reliance Electric and Engineering Company, Cleveland, Ohio): Prof. Koenig's papers represent an important contribution to the analysis of rotating machinery. This general type of approach, toward which Gabriel Kron has contributed so much, gives the engineer a powerful tool. The importance of this tool is increasing with the development and use of digital computers and with the continuing efforts to reduce the physical size of electric machines. Prof. Koenig's paper adds impetus to this general movement.

---

* An exception exists when the coils are not wound in the same direction and are not on a common magnetic circuit, a condition quite unlikely to exist in machines where every attempt is made to maintain symmetry.

It would probably be easier for an induction machinery designer to follow the paper if the author had used notation and concepts more common to the literature of induction machinery. I changed some of the equations to the notation suggested by the AIEE Working Group on Letter Symbols for Induction Motors (this report was not available when Prof. Koenig wrote his paper) as this notation is more representative of the viewpoint of the machinery designer. The difference is more than just using different letters for the same quantities; rather, there are important differences in definition of just what the parameters are that define a machine.

The most important difference, perhaps, is that Prof. Koenig thinks of open-circuit reactance as a single quantity which he calls ωL_1. The machinery designer thinks of the open-circuit reactance as the sum of two components of reactance: leakage and magnetizing. Leakage reactance and magnetizing reactance are kept separate in the machinery designer's mind, and there are good reasons for it:

1. Leakage reactance and magnetizing reactance are calculated separately as two distinct quantities.
2. Torques of the motor depend almost entirely upon leakage reactance, not upon total reactance.
3. Leakage reactance is usually small compared to magnetizing reactance.

Table I is a comparison of the symbols used by Prof. Koenig with those proposed by the AIEE Working Group on Letter Symbols for Induction Motors. At first thought, it might seem that dividing open-circuit reactance into two components (leakage and magnetizing) would lead to clumsier equations. Prof. Koenig's equations 1(A) and (B) of part I, rewritten in induction-motor letter symbols, become

$$V_1=[r_1+j(x_1+x_M)]I_1-jx_M\frac{I_{22}}{b} \qquad (1A')$$

$$-V_2=\left[-j\frac{x_M}{b^2}\right]bI_1+\left[r_{22}+j\left(x_{22}+\frac{x_M}{b^2}\right)\right]I_{22} \qquad (1B')$$

These two equations are still not so familiar to the machinery designer because the secondary quantities are given in secondary terms; if they are both rewritten in primary terms they become

$$V_1=[r_1+j(x_1+x_M)]I_1-jx_MI_2 \qquad (1A'')$$

$$-bV_2=-jx_MI_1+[r_2+j(x_2+x_M)]I_2 \qquad (1B'')$$

or, in matrix form, using impedances

$$\begin{bmatrix} V_1 \\ -bV_2 \end{bmatrix}=\begin{bmatrix} z_1+z_M & -z_M \\ -z_M & z_2+z_M \end{bmatrix}\begin{bmatrix} I_1 \\ I_2 \end{bmatrix}$$

This is recognizable instantly as the mesh-circuit matrix equation for Fig. 2. Thus, part of Prof. Koenig's "transformation of variable" converted secondary quantities to primary terms; the other part converted from mesh to branch form. Both conversions were made by the single transformation of variable.

Table I. Comparison of Symbols

Quantity	Koenig Symbol	Proposed Standard
Primary reactance, total	ωL_1	$x_1+x_M=X_s$
Ratio of primary turns to secondary turns	n	b
Mutual reactance	ωM_{12}	
Magnetizing reactance	ωnM_{12}	x_M
Secondary reactance, total, referred to itself	ωL_2	$x_{22}=\frac{x_M}{b^2}$
Rotor resistance, referred to itself	r_2	r_2/n^2
Open-circuit primary impedance	Z_1	z_1+z_M
Mutual impedance	Z_{12}	z_M
Leakage impedance of primary	Z_1-nZ_{12}	z_1
Leakage impedance of secondary, referred to primary	$n^2Z_2-nZ_{12}$	z_2
Rotor resistance, referred to primary	$R_2=n^2r_2$	r_2
Secondary current, actual	I_2	I_{22}
Secondary current, referred to primary	I_b	$I_2=I_{22}/b$

The branch-circuit matrix equation for the circuit of Fig. 2 can be written, in induction motor letter symbols, by inspection, as follows

$$\begin{bmatrix} V_1 \\ V_1-bV_2 \end{bmatrix}=\begin{bmatrix} z_1+z_M & z_1 \\ z_1 & z_1+z_2 \end{bmatrix}\begin{bmatrix} I_a \\ I_b \end{bmatrix} \qquad (7')$$

This is Prof. Koenig's equation 7, except in induction-motor letter symbols. Note that three of the four elements in the impedance matrix are actually simpler in induction-motor nomenclature than in Koenig's. In this case, at least, splitting the open-circuit primary reactance into two elements actually led to fewer, rather than more, terms in the elements of the matrix. So much of the performance is a function of leakage impedances, that time is saved by using leakage impedances.

A word of caution is in order. Magnetizing reactance x_M is often loosely referred to as "mutual" reactance. In ordinary induction-motor theory, magnetizing and mutual are the same only when the turns ratio is unity. Moreover, in the more general case as typified by Fig. 3 and equation 10, the mutual reactance may be quite different from magnetizing reactance. However, the M's in equation 10 are all functions of air-gap ("magnetizing") flux, even though they are multiplied by cosines of angular positions.

One purpose of this discussion was to show the correlation between Prof. Koenig's equations and conventional induction-motor equations. A second purpose is to suggest that it may be worth while to think of the total self inductance of a coil as consisting of the sum of two components:

1. Inductance due to "leakage" fluxes.
2. Inductance due to air-gap fluxes.

(The latter is not necessarily mutual.)

Such a procedure might well have made unnecessary some of Prof. Koenig's later partitioning of a matrix so as to bring out leakage reactance directly. The same idea may be as useful in transient as in steady-state analysis.

C. Concordia (General Electric Company, Schenectady, N. Y.): The author states that one of his objectives is to develop a theory of synchronous and asynchronous machines based upon circuit theory. We agree that this is a very worthwhile objective, but do not agree with the very strong implication that the application of circuit theory to rotating machinery is new. To the contrary, the development of a circuit theory of synchronous machines was precisely the purpose and accomplishment of R. H. Park.[1] Engineers concerned with the application of synchronous machines have been using and extending such a circuit theory for over 30 years. Several other early references could be cited, but most of them are included in lists given in reference 2.

The five steps listed in the author's summary of part I appear to be indistinguishable from those followed in many of these references and we should like to inquire in what way the author's method differs from these.

REFERENCES

1. TWO-REACTION THEORY OF SYNCHRONOUS MACHINES. GENERALIZED METHOD OF ANALYSIS—PART I, R. H. Park. *AIEE Transactions*, vol. 48, July 1929, pp. 716–30.

2. SYNCHRONOUS MACHINES—THEORY AND PERFORMANCE (book), Charles Concordia. John Wiley and Sons, Inc., New York, N. Y., 1951 (see especially pp. 216–19).

Karl L. Wildes (Massachusetts Institute of Technology, Cambridge, Mass.): Prof. Koenig is to be commended for the clarity with which he has described his circuits and machines. In his words "a good map is essential," and an important element of a good map is a precise designation of currents and voltages as assumed for the problem at hand. Coil polarity markings are necessary, as are current and voltage directions. An arrow to represent current direction is clear and universally accepted. Prof. Koenig uses a voltage-direction designation which is not so widely used, but which is crystal clear and does not require the distinction, used by some writers, between E for voltage rise and V for voltage drop. The plus sign on the circuit diagram tells the whole story.

In the section *The 2-Winding Transformer* in part I, the author says, "Unfortunately there is no apparent direct method available for deriving the transformation used in this and the following examples." There are two criteria which can be used effectively to suggest effective transformations. The first is to observe whether the network matrix is unsymmetrical, as it often is, and if so, to ask what transformation will make the matrix symmetrical or diagonal. The second criterion often comes out of the physics of the problem. To illustrate this second criterion, consider Prof. Koenig's 2-winding transformer. The actual instantaneous currents are i_1 and i_2. The differential equations are, following Fig. 1

$$v_1=(r_1+L_1p)i_1-M_{12}pi_2$$

$$-v_2=-M_{12}pi_1+(r_2+L_2p)i_2$$

All of the inductances are variable because of saturation effects. The physics of the problem suggests that at least some of the inductance coefficients can be made more nearly constant if their associated fluxes can be free from the iron path. If the inductance portions of the above equations are written in integrated form, the linkages of the primary and secondary coils are

$$\lambda_1 = L_1 i_1 - M_{12} i_2$$

$$-\lambda_2 = -M_{12} i_1 + L_2 i_2$$

The fluxes are

$$\varphi_1 = \frac{\lambda_1}{N_1} = \frac{L_1}{N_1} i_1 - M_{12} \frac{i_2}{N_1}$$

$$-\varphi_2 = -\frac{\lambda_2}{N_2} = \frac{-M_{12}}{N_2} i_1 + L_2 \frac{i_2}{N_2}$$

The leakage flux, or flux chiefly in air, is

$$\varphi_1 - \varphi_2 = \frac{L_1}{N_1} i_1 - M_{12} \frac{i_2}{N_1} - \frac{M_{12}}{N_2} i_1 + L_2 \frac{i_2}{N_2}$$

$$= \frac{1}{N_1}[(L_1 - nM_{12}) i_1 + (nL_2 - M_{12}) i_2]$$

in which $n = N_1/N_2$.

In order to use a single mutual inductance term in both parentheses, the usual turns-ratio transformations suggest themselves as follows:

$$N_1(\varphi_1 - \varphi_2) = (L_1 - nM_{12}) i_1 + (n^2 L_2 - nM_{12}) \frac{i_2}{n}$$

Thus the linearization of inductance coefficients leads to the current transformation

$$[i_1'] = \left[\begin{array}{c} i_1' \\ \hline i_2' \end{array}\right] = \begin{bmatrix} 1 & \\ & 1/n \end{bmatrix} \left[\begin{array}{c} i_1 \\ \hline i_2 \end{array}\right] = [C^{-1}]\,[i]$$

For invariance of power

$$[v'] = \left[\begin{array}{c} v_1' \\ \hline v_2' \end{array}\right] = \begin{bmatrix} 1 & \\ & n \end{bmatrix} \left[\begin{array}{c} v_1 \\ \hline v_2 \end{array}\right] = [C]\,[v]$$

The operational impedance transformation is

$$[\boldsymbol{Z}'] = [C]\,[\boldsymbol{Z}]\,[C] = \begin{bmatrix} 1 & \\ & n \end{bmatrix} \begin{bmatrix} r_1 + L_1 p & M_{12} p \\ -M_{12} p & r_2 + L_2 p \end{bmatrix} \times$$

$$\begin{bmatrix} 1 & \\ & n \end{bmatrix} = \begin{bmatrix} r_1 + L_1 p & nM_{12} p \\ -nM_{12} p & n^2(r_2 + L_2 p) \end{bmatrix}$$

which gives the usual equivalent circuit of Fig. 2.

A more complicated example is the 2-phase induction motor as worked out by the tensor methods of Kron.[1] The fundamental four-by-four matrix of the machine is transformed to positive- and negative-sequence right-angle components and then the resulting matrix

$$[\boldsymbol{Z}] = \begin{array}{c|c|c|c|c|} & s1 & s2 & r1 & r2 \\ \hline s1 & R_s + jX_s & & jX_m & \\ \hline s2 & & R_s + jX_s & & jX_m \\ \hline r1 & jX_m(1-v) & & R_r + jX_r(1-v) & \\ \hline r2 & & jX_m(1+v) & & R_r + jX_r(1+v) \\ \hline \end{array}$$

immediately suggests that symmetry can be achieved through multiplication on the left by

$$[C] = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \dfrac{1}{1-v} & \\ & & & \dfrac{1}{1+v} \end{bmatrix}$$

giving an equivalent circuit for positive-sequence quantities and another for negative-sequence quantities with no interference between them.

A second equivalent circuit can be found through ignoring the rotor axes of $[\boldsymbol{Z}]$ resulting in a two-by-two diagonal matrix. The symmetrical or diagonal nature of the matrix does not reveal anything about the nonlinearity of the reactances. This must come from a leakage-flux consideration as in the transformer example.

In addition to its power to suggest equivalent circuits, the matrix method is very effective in organizing the various points of view in a circuit problem. For example in three-phase machine analysis, the transformations among the several sets of coordinates (a, b, c; $0, 1, 2$; $0, \alpha, \beta$; $0, d, q$; and $0, f, b$) can be nicely manipulated and the corresponding sets of self- and mutual impedances can be kept clearly in view by means of the matrix organization. It should be noted that the transformations of equation 38 of part I are the Clarke transformations which change the a-b-c axes into 0-α-β axes and which are in turn easily related to the 0-d-q axes.

REFERENCE

1. TENSOR ANALYSIS OF NETWORKS (book), Gabriel Kron. John Wiley and Sons, New York, N. Y., 1939.

H. E. Koenig: The comments of the discussors are greatly appreciated.

As Mr. Saunders has indicated, the five idealizations stated in *Assumptions, Reference Systems, and Definitions* are made primarily for the sake of simplicity and need not be made in all studies. For example, if only the steady-state solution is required for balanced conditions of operation all harmonics (including slot harmonics), the saturation effects can be included to the extent that they are included in the classical procedures. In fact, the method of analysis as presented here reduces to the classical method of analysis when the restricted problem is considered.

Mr. Saunders' discussion on the sign of the mutual inductance, M_{jk}, and the voltage of mutual induction, $(d/dt)(M_{jk} i_k)$ brings up a very important point. When the rotating machine is considered as a dynamic network, M_{jk} is a function of rotor position and must take on both positive and negative values. Hence, the magnetic polarization marks must be considered as reference marks. To say they indicate the actual magnetic polarization or to say they are so located as to make $M_{jk} > 0$ is meaningless. Since the magnetic polarization marks are to be considered as reference marks, their location with respect to the terminals of the coils are completely arbitrary and one logically chooses their locations in such a way as to produce maximum useful symmetry in the analysis. It is believed that the useful symmetry in the functions describing the variation in the M_{jk} terms with rotor position, and the symmetry in the literal form of the differential equations obtained by placing all the magnetic reference marks at the tail (or head) of the current reference arrows is far more valuable than any advantage obtained by attempting to choose the markings so as to make all M_{jk} terms positive at zero rotor position. If the coils are all fixed with respect to each other, it is necessary to use more than one set of polarization marks when the coils are not on a common magnetic circuit and when one insists that all the M_{jk} terms are to be represented by positive numbers. When considered as references, one set of markings is sufficient for all coils. Here again, it is believed that the simplification and useful symmetry obtained by choosing all the magnetic reference marks at the tail (or head) of the current references is more valuable than the advantages obtained by choosing the polarity markings so as to make all M_{jk} terms positive numbers. Finally, when the polarization marks are considered as references and the M_{jk} terms allowed to take on positive and negative values like the currents and voltages, a much extended class of problems can be studied by a single procedure. This is the sort of thing which leads to increased teaching efficiency and organization in thinking regarding the equations of machines.

Another way to determine whether M_{jk} is to be represented by a positive or negative number at any instant of time is to observe the polarity of the voltages at the terminals of the jth and kth coil. If at any instant, the polarities of the voltages of the two coils are the same with respect to the magnetic polarization references, then M_{jk} is represented by a positive number at that instant. If they are opposite, then M_{jk} is represented by a negative number.

Mr. Veinott has very clearly pointed out the similarity between the notation and concepts used by the induction motor designer and those used in this paper. The leakage-reactance concepts (differences between self- and mutual inductances) are most valuable and convenient in the study of electric machinery. I only wish to clarify one point in connection with his discussion. In the study of the 2-winding transformer and the induction motor one can easily start with the leakage-reactance or leakage-inductance concepts. However, the point of the presentation given in the paper was to show how these concepts can be obtained as a mathematical transformation of variable. In the study of a 4-winding transformer, for example, it is difficult to start with a set of meaningful leakage-reactance concepts. By following the mathematical transformation of variable as outlined for the 2-winding transformer, the difference terms (differences between self- and mutual inductances) appear automatically, and one can pursue the analysis with complete assurance.

In regard to Mr. Concordia's discussion, it is sufficient to say that it is stated in the introduction to part I that, "The idea that the rotating machine can be considered as a network with time-varying coefficients is not new, and many of the results arrived at in this paper have been arrived at by previous authors using many varied methods of analysis." The references sighted were intended to be only a sample of the vast quantity of material relating to this area of study.

It is further believed that there is a substantial difference between the basic philosophy and the organization of the problem as presented in this paper and that presented in the references sighted by Mr. Concordia. Nowhere in the material presented in this paper does one encounter the troublesome concepts of primitive networks, reference frames, and tensor mathematics.

I wish to thank Mr. Wildes for his discussion in which he points out certain criteria and studies which lead to a useful transformation of variable. Obviously, the statement made in the paper concerning the derivation of these transformations was not clearly presented. When speaking of a mathematical derivation, one usually implies that definite mathematical processes are available for arriving at a unique answer. It is this type of mathematical derivation the author had in mind when making the statement, "Unfortunately, there are no apparent direct methods available for deriving the transformations used in this and the following example." As Mr. Wildes clearly points out, the transformations sometimes come from a study of the form of the equations or from a study of the physical device which the mathematics describes.

Significant Measurements for Determining the Stability of High-Temperature Magnet-Wire Insulation

A. L. SCHEIDELER
ASSOCIATE MEMBER AIEE

Synopsis: Test methods for evaluating the stability of high-temperature magnet-wire insulation are described. Reduction in dielectric strength during uninterrupted exposure to high temperature is used to define the insulation life which would be obtained with comparatively ideal operating conditions. Moisture resistance, as measured by power factor after humidification, is proposed as a criterion for determining the ability of a material to withstand such adverse operating hazards as moisture absorption or mechanical shock. Results have been obtained which indicate that deterioration rates, calculated from measured reduction in dielectric properties at high temperatures, may be experimentally verified at lower temperatures. When this is true, justification is provided for the extrapolation of data to lower, operating temperatures. Conversely, the absence of similar aging trends at different, accelerated temperatures should serve notice that care must be taken in the extrapolation of data to low temperatures.

THE selection of an adequate insulation system is a major problem for the designer of high-temperature electric apparatus. Many materials are available which will meet initial design requirements and their suitability in this regard can be determined by comparatively simple and routine tests. Assessing the ability of such materials to maintain adequate properties throughout desired equipment life is, however, somewhat more complicated. Unfortunately the chemist's pace in combining the elements to form new materials has far exceeded his progress in evaluating their suitability for application by the designer. In most cases it is necessary to evaluate experimentally the aging characteristics of any material which looks promising in the light of initial tests.

The primary objective of most such evaluation programs is to determine the stability of composite materials under the variety of conditions encountered in service. Accordingly, much emphasis in recent years has been placed upon the "functional" evaluation of materials and apparatus. However, any developmental program for the evaluation of such material is quite costly in time, labor, and equipment. It is highly desirable, then, to conduct the tests so that something may also be learned of the deterioration processes. Such information will aid in the understanding of the aging mechanisms and provide more useful data for the design of new insulating systems.

Factors Influencing Aging Procedures

The institution of any realistic aging program requires first that the pertinent factors affecting deterioration be established. A second and much more difficult task is to determine the degree to which each of these should be considered. The aging factors may differ with each type of apparatus and even within apparatus groups. For example, the problem of exposure to ambient moisture is much more significant in a ventilated dry-type transformer than in a sealed unit.

The primary aging factor for magnet-wire insulation in dry-type transformers is heat. However, the material in a transformer may at intervals be subject to mechanical shocks (short circuits and switching surges) and to moisture (high humidity ambient). These are not, in general, primary sources of deterioration, but the insulation must be able to withstand any deleterious effects caused by moisture and mechanical shock.

It is difficult, if not impossible, to establish typical test conditions which will produce results applicable, even on a relative basis, to all reasonable service usages. A more practical and informative approach is to bracket the typical operating cycles with the test procedures, thus per-

Fig. 1. Test piece used in evaluating the thermal stability of high-temperature magnet-wire insulation. Data presented are for an acrylic ester resin-impregnated asbestos - clay sheet covered with glass yarn. Final treating varnish for data presented in Fig. 2—phenolic A; for all other data—phenolic B

Paper 54-71, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 19, 1953; made available for printing November 23, 1953.

A. L. SCHEIDELER is with the General Electric Company, Pittsfield, Mass.

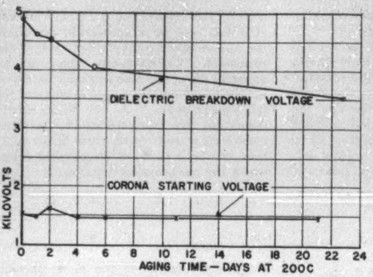


Fig. 2. Relationship between corona starting voltage at 25 C and hot dielectric strength after exposure to thermal aging at 200 C

mitting interpolation of results to any desired conditions. The scope of such a program must, of course, be maintained within the bounds of economic practicability. This bracketing of field conditions may be accomplished effectively by evaluating a composite magnet-wire insulation: 1. under comparatively ideal conditions (minimum of thermal cycling, moisture exposure, mechanical shock), and 2. under circumstances which allow performance measurement in an adverse situation. Material life at these two extremes may be determined as maximum and minimum life.[1]

Aging Criteria

The selection of a criterion for the measurement of aging rates and insulation life is an important one. Physical, chemical, and electrical tests have all been used by investigators as measures of insulation degradation. Useful information may be obtained by the proper interpretation of the results of such measurements. However, the final criterion for the usefulness of a given high-temperature insulation in electric apparatus is its ability to withstand electric stress under service conditions. All other aging factors are most significant in so far as they affect the ability of the insulation to maintain adequate dielectric properties.

It is also highly desirable, if possible, to utilize a test for evaluating aging characteristics which is in itself nondestructive. A test which does not destroy the sample may be repeated as a function of aging time and the aging trends determined on an individual sample. Destructive testing, on the other hand, may be accomplished only once on a given test piece. Many more test pieces are required to establish the relationship between insulation property and aging time.

Test Piece

The evaluation of high-temperature magnet-wire insulation for dry-type transformers requires that the factors just mentioned be considered. Careful attention should also be given to the design of the test piece. The most realistic results will be obtained when the material is tested in a configuration where the physical and electric stresses are representative of those found in actual apparatus.

The sample structure illustrated in Fig. 1 has proved very successful in the evaluation of materials applied to comparatively large rectangular wire such as is used in the larger dry-type transformers.[2] Any combination of insulations to be investigated is applied to the wire using standard factory processes. The insulated wire is then cut into 10-inch lengths. Two pieces are placed in a jig, spread apart at the ends, and the middle 6 inches tied together with glass yarn by an experienced operator. The completed sample is then given a standard treatment with the desired treating varnish.

When completed, the test piece consists of two insulated wires which have been exposed to the mechanical and thermal stresses normally encountered in the manufacturing operations. Since the wires are placed in a typical turn-to-turn relationship, realistic dielectric tests may be applied to the insulations using the copper wires as electrodes. Furthermore, the sample is small enough to represent a small section of a transformer winding with negligible thermal gradient. Thermal aging may be accomplished realistically in constant temperature ovens.

Maximum Life

The maximum life may be expected from a high-temperature magnet-wire insulating structure when it is not subjected to an undue amount of moisture or mechanical abuse. Two nondestructive tests have been considered for use in measuring deterioration rates of the sample under these comparatively ideal conditions. These are: 1. power-factor measurements on a dry sample at room temperature, and 2. corona starting voltage measurements[3] under similar conditions. The effect of aging time at elevated temperatures on these properties and on the hot dielectric strength of similar samples are illustrated in Figs. 2, 3, and 4. No significant correlation has been noted between the corona starting voltage and the dry dielectric strength of the material. A correlation between dry power factor and dielectric strength may exist. However, it has been found very difficult to avoid the influence of ambient humidity conditions on the dry power-factor measurements. (Even minutes of exposure to atmospheric humidity significantly affect results after the material has deteriorated to some extent.) This, combined with the very small effects of aging, lead to the rejection of dry power-factor measurements as means of establishing maximum life characteristics.

The reduction of dry dielectric strength as a function of aging time at elevated temperatures, Fig. 4, has been utilized as a means of estimating a reasonable maximum life for the material. The significance of dielectric strength criteria has

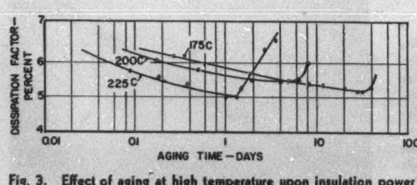


Fig. 3. Effect of aging at high temperature upon insulation power factor in a dry atmosphere at 25 C

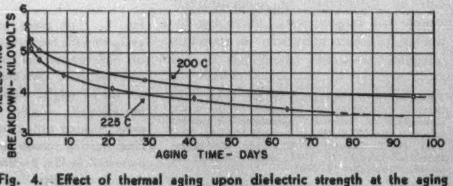


Fig. 4. Effect of thermal aging upon dielectric strength at the aging temperature

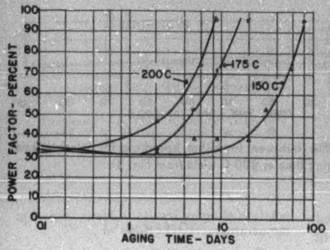


Fig. 5 (left). Thermal aging characteristics as determined by periodic power-factor measurements after 24 hours exposure to 90-per-cent relative humidity, 90 degrees Fahrenheit

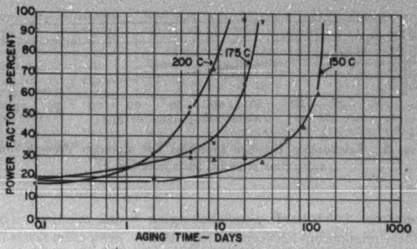


Fig. 6. Thermal aging characteristics as determined by periodic power-factor measurements after 4 hours exposure to 90-per-cent relative humidity, 90 degrees Fahrenheit

been well established in other investigations[4,5] and since samples are small and relatively simple to prepare it is feasible to make a large number of tests. Dielectric strength may then be determined as a function of aging time.

Minimum Life

The minimum life of materials used in ventilated dry-type transformers may, in general, be determined by the moisture-resisting properties of the insulation. Dielectric and mechanical stresses normally encountered in service are comparatively minor sources of insulation degradation. However, when the material has deteriorated due to thermal aging by flaking, crazing, or charring, it no longer provides the mechanical bonding or the moisture resistance necessary for operation under adverse conditions.

Although dielectric breakdown after humidification should serve as an adequate criterion of minimum life, nondestructive tests are available. Both insulation resistance and power-factor measurements after humidification on an aging sample have been considered. The use of 60-cycle power factor measurements were selected as superior since they are made at the operating frequency, effects of charging current are eliminated, and the power-factor measurement is more accurate in practice. The test procedure given in the following has proved very useful.

At appropriate intervals, the samples are removed from the aging ovens, cooled to room temperature, and placed in a humidity chamber (90-per-cent relative humidity, 90 degrees Fahrenheit). After 24 hours exposure to high humidity, power-factor measurements are made and the samples removed from the humidifier. They are allowed to dry for 24 hours at room conditions and then replaced in the oven. The cycle is repeated at intervals until the power factor, after exposure to high humidity, reaches a value of 96 per cent. This is taken as the relative end of life when using power-factor measurements as a criterion. Typical results for class-*B* magnet-wire insulations are illustrated in Fig. 5.

Measurements were made on a group of class-*B* samples in a manner similar to that described. However, they were humidified for 4 hours instead of the usual 24. Results similar to those of the 24-hour humidified samples were obtained, Fig. 6, with the exception that the time to the relative end of life has been extended. The 24-hour test period has been selected as most practical since small variations in the humidification time have less effect upon results.

Sealed units are, with the exception of limited exposures during maintenance, free from the deleterious effects of moisture. Minimum life under such conditions will most probably be due to reduced mechanical strength of the winding caused by thermal aging of the bonding varnishes. The structure will then be particularly vulnerable to short circuits and switching surges. Varnish decomposition during aging in the sealed atmosphere, which may easily be measured by a periodic determination of moisture resistance, should be useful in interpreting the minimum life of magnet-wire insulation in sealed as well as ventilated units.

Mechanisms of Insulation Deterioration

The ability to define the process by which an insulation under study approaches failure would be very useful in understanding more clearly the mechanisms of insulation deterioration. Recently, Dakin[6] and Malmlöw[7] have advocated that thermal deterioration of insulation be treated as a chemical reaction which, as a function of temperature, obeys the well known Arrhenius equation. Dakin then demonstrated that, in general, the logarithm of insulation life is a linear function of the reciprocal of the absolute aging temperature. Such a relationship is useful in extrapolating to lower temperatures the results obtained by accelerated aging tests at high temperatures.

Malmlöw has effectively extended the line of reasoning used by Dakin in studies of the aging characteristics of cellulose. Certain mechanical properties (for example, tensile strength, bursting strength) have been shown to deteriorate in a manner which indicates the presence of two chemical reactions. The effects on the properties, due to these reactions, may be separated, allowing a much more comprehensive analysis of the aging characteristics of a material.

It is interesting to note that similar effects are encountered in typical high-temperature magnet-wire insulation using dielectric measurements as a criterion. The ordinate of Fig. 6 (per cent power factor $=100 \sin \delta$) has been replaced by ln dissipation factor ($\ln \tan \delta$) and the data replotted in Fig. 7. A linear relationship between ln dissipation factor and aging time is apparent at each aging tem-

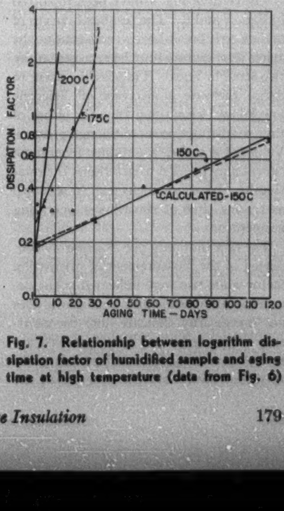


Fig. 7. Relationship between logarithm dissipation factor of humidified sample and aging time at high temperature (data from Fig. 6)

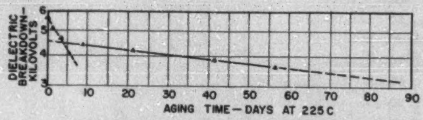


Fig. 8. Relationship between logarithm hot dielectric strength and aging time at 225 C (data from Fig. 4)

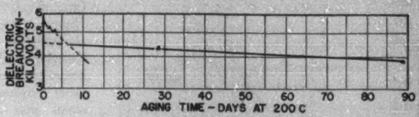


Fig. 9. Relationship between logarithm hot dielectric strength and aging time at 200 C (data from Fig. 4)

perature. It has been shown[8] that such a linear relationship between the logarithm of the pertinent physical property (in this case dissipation factor after humidification) and aging time is indicative of a first-order chemical reaction. The relationships between the pertinent variables for such a reaction have been developed in the Appendix and are described by equation 1

$$\log D = B - At\epsilon^{-E/RT} \qquad (1)$$

where

D = dissipation factor after humidification
A = total frequency of encounter between reactant molecules
E = "energy of activation"
R = molar gas constant
T = absolute temperature
B = the level of D at zero time

When such definable relationships occur, the data obtained at accelerated aging temperatures may be used to calculate A and E/R. These values may, in turn, be used to determine the relationship between dissipation factor after humidification, and aging time at some lower operating temperature. Using the form of equation 1 values of A and E/R were obtained from the data, Fig. 7, at 200 and 175 degrees centigrade (C). The calculated 150 C curve obtained from these values agrees very well with the experimental 150 C curve.

The data of Fig. 4 have been replotted in Figs. 8 and 9. The resultant curves of ln dielectric breakdown versus time at the aging temperature may be represented by two straight lines. It is possible, then, to define the rate of decrease of dielectric breakdown with aging time for either of the two distinct parts of the aging curve. Values of A and E/R obtained from these curves may be used to determine the relationship between dielectric breakdown and aging time at some lower operating temperature.

The fact that only one reaction was observed in the measurement of dissipation factor while two were indicated in the determination of dielectric breakdown does not necessarily indicate any inconsistency. The orders of time involved are such that the dissipation factor is probably indicating only the progress of the first reaction. The values of humidified dissipation factor are too high to measure during the extended aging times, in which the second reaction, indicated by dielectric breakdown, would predominate.

The amenability of the data just presented to analysis in terms of chemical reactions indicates that physical measurements hold promise as a means of interpreting the chemical phenomena influencing the deterioration of insulating structures. Such analyses may also serve as a useful tool in the functional evaluation of insulation. The fact that curves calculated from measurements made at accelerated temperatures may be experimentally verified at lower temperatures indicates that the phenomenon of deterioration is the same throughout the range considered. Justification is thereby provided for the extrapolation of such results to operating temperatures.

Conversely, the absence of similar trends at different accelerated temperatures should serve notice that care must be taken in the extrapolation of data to lower temperatures. Such a situation may arise due to the influence of temperature on curing characteristics or upon the actual deterioration mechanisms.

Conclusions

1. A test method has been developed for the evaluation of high-temperature magnet-wire insulation which effectively brackets typical service conditions. Consideration may then be given to a variety of service requirements in the interpretation of aging results.

2. Reduction in dielectric strength at the aging temperature may be used effectively to determine maximum life characteristics of magnet-wire insulation.

3. The relative moisture resistance of magnet-wire insulation as determined by dielectric loss measurements after exposure to high humidity may be interpreted to determine the minimum life of a magnet-wire insulation.

4. Test results which illustrate the manner of insulation deterioration may be useful in interpreting the deterioration mechanisms.

5. Determination of deterioration rates for a given material may show that the phenomenon of deterioration is the same at widely different accelerated aging temperatures. Justification is thereby provided for extrapolation of aging results to lower operating temperatures.

6. Caution should be observed in extrapolating to lower temperatures data which do not indicate the presence of the same deterioration mechanisms at different accelerated aging temperatures.

Appendix I. Reaction Rates and Physical Deterioration

If it may be assumed that one chemical reaction is predominant in influencing insulation deterioration, we may write

$$F(C) = -Kt \qquad (2)$$

where $F(C)$ is a function of the concentration of the most important chemical constituent and K is a rate constant dependent on temperature but not on time t. In order to utilize this relationship in analyzing the aging characteristics of a given material, it must be presumed that some pertinent physical property (i.e., tensile strength, electric strength, dissipation factor) is a function of the concentration C. Experimental verification will, of course, be necessary to substantiate this hypothesis.

If the relationship between concentration and physical property is valid, we may write

$$F(D) = -Kt \qquad (3)$$

where $F(D)$ is a function of the physical property D under consideration.

The reaction rate constant K has been proposed by Arrhenius to be of the form

$$K = A\epsilon^{-E/RT} \qquad (4)$$

The validity of this form for representing many reactions has been proved by time. In equation 4

A = total frequency of encounter between reactant molecules
E = "energy of activation"
R = molar gas constant
T = absolute temperature

Substituting in equation 3 from equation 4 we have

$$F(D) = -At\epsilon^{-E/RT} \qquad (5)$$

It can be seen from equation 5 that if the form of $F(D)$ is known and if A and E/R can be evaluated, the relationships between D, T, and t are defined. For example, if D represents dielectric strength, then the dielectric strength of a material at time t and temperature T can be determined. By defining $F(D)$ from data obtained at high temperatures, the performance of a material at lower operating temperatures can be predicted by means of relatively short-time tests.

References

1. Organo-Silicone Compounds for Insulating Electric Machines, **T. A. Kauppi, G. L. Moses.** *AIEE Transactions (Electrical Engineering)*, vol. 64, March 1945, pp. 90–93.

2. "Terratex"—A Thin Flexible Inorganic Insulation, **T. R. Walters.** *AIEE Transactions*, vol. 67, pt. I, 1948, pp.123–27.

3. Development of Corona Measurements and Their Relation to the Dielectric Strength of Capacitors, **R. J. Hopkins, T. R. Walters, M. E. Scoville.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1643–51.

4. Aging Characteristics of Dry-Type Transformer Insulation at High Temperature, **H. C. Stewart, L. C. Whitman.** *AIEE Transactions*, vol. 67, pt. II, 1948, pp. 1600–07.

5. Aging Evaluation of Dry-Type Transformer Insulating Systems, **H. C. Stewart, L. C. Whitman, A. L. Scheideler.** *AIEE Transactions*, vol. 72, pt. III, April 1953, pp. 267–77.

6. Electrical Insulation Deterioration Treated as a Chemical Rate Phenomenon, **Thomas W. Dakin.** *AIEE Transactions*, vol. 67, pt. I, 1948, pp. 113–22.

7. Thermal Aging Properties of Cellulose Insulation Materials, **G. Malmlöw.** *Transactions of the Royal Institute of Technology*, Stockholm, Sweden, no. 19, 1948.

◆

Discussion

Paul Narbut (Westinghouse Electric Corporation, Sharon, Pa.): It may be of interest to consider the practical significance of the results of the tests presented in the paper, particularly, to estimate how closely the author came to specifying a meaningful test for evaluation of life of turn insulation in dry-type power transformers.

I am substantially in agreement with the author in regard to some of his general statements, for instance, that the final criterion for the usefulness of electrical insulation should be its ability to withstand electrical stresses under service conditions; also that the most realistic test results will be obtained when the material is tested in configuration where the physical and electrical stresses are representative of those found in actual apparatus. These statements express the functional approach to insulation evaluation. Insulation should be considered in good condition if it is adequate to perform the intended function.

To evaluate the author's conclusion regarding the validity of his tests, I should like to give results of another test, which was made on disk-type coils, typical of high-voltage winding construction, using shop standard glass-varnish conductor insulation. The coils were wound of double conductor, which permitted tests on the turn insulation between the two strands. The area tested was large. This condition, incidentally, was not met in the author's tests. The coils were mounted and braced in a standard manner and were given a shop standard varnish treatment, normally received by dry-type power transformers with class *B* insulation. Thus, the form of the samples represented realistically the actual application, as well as the shop practices.

Dielectric tests on freshly prepared samples in dry condition gave a withstand value of 1,200 volts on a 1-minute 60-cycle test between turns. The samples were then subjected to several temperature-aging cycles at an average of 260 degrees centigrade (C), with a total duration of over 20 days. The degree of varnish deterioration resulting from this aging was quite obvious, and was severe. After aging, the samples were tested again, this time in the atmosphere of 92 per cent humidity. The 1-minute withstand voltage under this condition was 1,100 volts.

The question arises: is the turn insulation on these samples still in good serviceable condition, or not? The dielectric test which, as we have agreed, represents one of the basic requirements, indicates little or no deterioration of insulation. ~~The loss factor~~ of the aged insulation, to be sure, is high but, frankly, we feel that this is unimportant as long as the insulation performs its intended function. The aged insulation has lost much of its mechanical strength, abrasion resistance, etc. But if these are the critical requirements, then a valid test to determine the end of insulation life should be based on these properties, not on arbitrarily specified values of dielectric strength, or loss factor. Referring to the question at the beginning of this paragraph, we do not have the answer. Unfortunately, the paper does not either.

It is of interest to ask the question: why is there such a discrepancy between results reported by Mr. Scheideler, and those reported in the foregoing, regarding the loss of dielectric strength of class *B* insulation upon aging. The discrepancy is still greater when we consider tests on class *B* insulation referred to in the paper.[1] According to the latter, the life of class *B* insulation (phenolic varnish-impregnated asbestos) is only approximately 11 days at 225 C. This is based on the reduction of the dielectric strength to one-half of the original value. If the same end point criterion is applied to the author's data on Fig. 4, it appears possible that this end point will never be reached. Certainly, it would not be reached in 100 days of aging at 225 C.

The reason for this difference is, of course, the role which is played by the organic impregnant in the sample. If the presence of this impregnant, in the form of a continuous film, contributes to the dielectric strength of the combined insulation, then a relatively rapid initial deterioration of the dielectric strength may be expected. As aging progresses, however, the dielectric strength will ultimately settle down to an asymptotic value, determined by the separation provided by the inorganic fiber. Fig. 4 strongly indicates that such asymptotic value of dielectric strength exists.

If the design of the apparatus is based on this, so to speak, base value of the dielectric strength, any prior decrease in dielectric strength on aging may not be considered a realistic criterion to define the end point of insulation life. This is so in spite of the fact that the test data obey the Arrhenius time-temperature dependence, the first order reaction law, etc. It is apparent that an approach different from that used by the author is required to find the answer as to what constitutes the end point of insulation life. In my opinion the only realistic approach is that based on a correct evaluation of the actual function performed by the insulation in the specific apparatus under consideration. It is a functional approach.

This discussion may appear somewhat critical of the author's efforts. It is true that I do not consider the problem solved of determining the life-temperature dependence for class *B* conductor insulation for dry-type power transformers. Nevertheless, I feel strongly that the data reported are a valuable contribution to the available information on insulation aging, and that the author should be commended for their presentation.

Reference

1. See reference 5 of the paper.

A. L. Scheideler: In answer to Dr. Narbut's question concerning the applicability of the test methods for determining an end point of insulation life, it would be well to consider the tests on the disk-type coils which he has described.

Since the discusser's dielectric test in a humid atmosphere reveals little or no decrease in electrical strength with aging, it seems reasonable to conclude that the material would be just as good, or better, when dry. Assuming that this relationship holds at the operating temperature, it may be said that no end point has been reached under maximum life conditions.

According to Dr. Narbut "the degree of varnish deterioration resulting from this aging was quite obvious and severe." The test just described has given little or no indication of these effects. One should be careful, however, in condemning the dielectric strength test in a humid atmosphere, as a means of measuring deterioration of the film. It must be remembered that stress was applied for 1 minute only. Since dielectric breakdown caused by varnish deterioration would probably be thermal in nature, a longer time test, approaching more closely the continuous stress conditions of service, may be desired.

At any rate, measurements of loss factor on such a test piece would show the progressive deterioration of the varnish film. An end point (96 per cent power factor) may be chosen to define the relative life for a given combination of materials. To place this figure on an absolute basis, it is of course necessary to compare such end points with those of combinations which have been demonstrated as satisfactory under service (truly functional) conditions. The figures are very useful in themselves in comparing the relative merits of different systems.

I have no particular prejudice against the use of mechanical tests in obtaining this kind of information. I know of no single test, however, which is as effective as the humidified loss factor measurement in synthesizing the effects, i.e., crazing, cracking, and chemical degradation, which may lead to mechanical failures and thermal breakdowns.

The emphasis which should be placed on these minimum life tests depends, as Dr. Narbut has indicated, upon the importance which is attached to the role of the organic impregnant in the test piece. It is obvious that gross deterioration of the varnish will reduce the mechanical strength of the complete coil structure and the hazards of short circuits and switching surges will be increased. The relatively free entrance of moisture will, if it does not cause actual breakdown, contribute to further deterioration of the insulating structure. Finally, in the absence of data to the contrary, it must be presumed that some hazard exists from the abrasion of glass fibers in the presence of minute vibrations. It would seem that operating temperatures, which result in gross deterioration of varnish films during a reasonable service life, should be avoided in power transformer design.

It might be admitted that a dielectric strength criterion is open to question if apparatus designs are dependent only on what Dr. Narbut calls the base value of the component insulation. The criterion becomes very realistic however if the structures are designed to maintain at least a given percentage of initial levels throughout the expected life of the equipment. Such an approach seems necessary if factory-proof tests are to provide any figure of merit for subsequent performance in the field.

I agree heartily that the only realistic approach to the problem of insulation evaluation must be based on consideration of the actual function performed by the material in specific apparatus. I would point out, however, that the only direct means of doing this is by evaluating apparatus performance in the field. Furthermore, it is not economically feasible in the design of large apparatus to evaluate all proposed materials as integral parts of a complete structure. Tests on components represent the only means by which engineering information can be obtained on large numbers of materials in a reasonable length of time. It seems logical that those tests should be chosen which provide the maximum amount of useful information.

Analysis of a Single-Core Magnetic Amplifier with Real Rectifier and Core Functions

MAX FRANK
ASSOCIATE MEMBER AIEE

SAUL RABOTNICK
NONMEMBER AIEE

J. R. WALKER
ASSOCIATE MEMBER AIEE

A REVIEW of the literature concerned with magnetic amplifiers has disclosed a myriad of articles dealing with the load circuit and sometimes control circuit response of single-core amplifiers employing the self-saturation principle.[1-4] Each of these discussions has utilized specific assumptions which are in each case aimed toward an ultimate simplification of the mathematical manipulations which finally lead to the required response. In many cases the predicted magnetic amplifier response suffers because of the assumptions thus used.

One of the most frequent assumptions employed by many authors is that of the existence of perfect rectifiers for use in the self-saturation circuit. This assumption limits the usefulness of most analyses in which it is utilized.

Other assumptions often employed include the following:

1. Magnetic material with zero area and infinite slope in the flux-current loop.
2. Magnetic material with zero area and finite slope in the flux-current loop.
3. Magnetic material whose dynamic loop is midway between the major a-c flux-current loop and the d-c loop.

Each of these assumptions is seen to specialize the consideration to an extent which may render the final conclusions inaccurate. With due respect to the types of analysis cited, it seems worth while to consider a rigorous mathematical analysis which is based upon the existence of finite forward and reverse rectifier resistances as well as a flux-current loop of finite area and slope.

Thus consider the fundamental single-core circuit shown in Fig. 1. Observe that a control circuit is not included with this core since in the present consideration the objective is to develop the load circuit relationships which will indicate the dependency of the load current on the rectification ratio and core constants.

Circuit Response

This objective will be accomplished by employing a linearized rectifier function and core function in appropriately chosen intervals which allow functional representation by defined constants. The solution of the governing differential equation of core flux shall be obtained for each interval of linear representation. Constants of integration for each interval solution shall be obtained from the boundary conditions. After the evaluation of the flux constants of integration has been completed, it will be possible to obtain the period-switching angles which define the boundaries of core and rectifier linearization.

Shown in Fig. 2 is the rectifier voltage-current function which is employed in the analysis. It is apparent from this figure that the back and forward resistances of the rectifier are assumed to be constant at R_B and R_F respectively.

The core magnetization function is shown in Fig. 3 with the appropriate cyclic intervals indicated thereon. A comparison of the assumed magnetization function with an actual flux-current-function is shown in Fig. 6. From these figures it may be observed a specialization has been imposed on the magnetization function to the extent that the width of the minor loops are assumed to remain unaltered at the magnitude given according to the major a-c loop and, further, that infinite slope exists when $\phi=\phi_s$ and $\phi=\phi_N$. According to measurements which have been made for higher grade core materials, it appears that such specializations are valid.

Since the sides of the magnetization function shown in Fig. 3 are assumed to be linear, it is possible to write the ampere-turn function

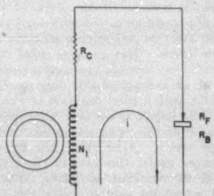


Fig. 1. Fundamental amplifier circuit

Paper 54-158, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted July 17, 1953; made available for printing December 7, 1953.

Max Frank, Saul Rabotnick, and J. R. Walker are with the Wayne Engineering Research Institute, Detroit, Mich.

List of Symbols

A_c = cross-sectional area of core, square centimeters
B_R = remanent flux density, gausses
B_s = saturation flux density, gausses
C_1 = core excitation constant, maxwells per second
C_2 = quality constant, second^{-1}
C_4 = quality constant, second^{-1}
E_M = a-c supply voltage, peak volts
h_0 = coercive force, ampere turns
i_1, i = inst. load current (general) amperes
i_{1-2} = inst. load current during time t_1-t_2
i_{2-3} = inst. load current during time t_2-t_3
i_{3-5} = inst. load current during time t_3-t_5
i_{5-6} = inst. load current during time t_5-t_6
i_{6-7} = inst. load current during time t_6-t_7
i_{7-8} = inst. load current during time t_7-t_8
k = core-stacking factor
l = path length of core, centimeters
N_1 = load circuit turns
Q = figure of merit or quality of the unsaturated circuit
R_1 = load resistance, ohms
R_c = load coil resistance, ohms
R_F = rectifier forward resistance, ohms
R_B = rectifier back resistance, ohms
R_{T_F} = total over-all load circuit resistance, forward
R_{T_B} = total over-all load circuit resistance, back
s = slope of flux-current function (ampere-turns per maxwell)
t_1, t_8 = beginning and end or closure time or the a-c cycle of consideration, seconds
t_2 = time end of negative loop knee, seconds
t_3 = time at initiation of core conduction or positive loop knee, seconds
t_4 = same as t_3 for materials with infinite slope
t_5 = end of core conduction, seconds
t_6 = end of rectifier positive conduction, seconds
t_7 = beginning of negative loop knee, seconds
ϕ = interval core flux, maxwells
ϕ_s = saturation flux, maxwells
ϕ_N = flux level from t_7 to t_2, maxwells
ϕ_{c1} = integration constant for period t_2 to t_3
ϕ_{c2} = integration constant for period t_5 to t_6
ϕ_{c3} = integration constant for period t_6 to t_7
ϕ_R = residual or remanent flux, maxwells
$g(\phi)$ = ampere-turn function

$$g(\phi)=N_1 i_1 = s\phi \pm h_0 \qquad (1)$$

Here the slope s of the loop sides is defined as $\frac{d}{d\phi}(g(\phi))=\frac{h_0}{\phi_R}$

The value of coercive ampere turns h_0 in actual magnitude must be obtained from flux-current loop data which is presented on the basis of sinusoidal core voltage excitation. The value of h_0 thus obtained is smaller than the normally published values contained in manufacturers' data because such data have been obtained from sinusoidal current excitation of the test core. It is well known that sinusoidal current excitation of a core will result in coercive current growth due to increased eddy current loss.[5] From the data obtained by sinusoidal voltage excitation of the core, the magnitudes of ϕ_R and ϕ_s may be obtained.[6]

The current becomes

$$i_1=\frac{s\phi}{N_1}\pm\frac{h_0}{N_1} \qquad (2)$$

If the applied voltage is given as $E_m \cos\omega t$, then the differential equation for the circuit shown in Fig. 1 becomes

$$E_m \cos\omega t = i_1 R_T + N_1 \frac{d\phi}{dt}\times 10^{-8} \text{ volts} \qquad (3)$$

Here the quantity R_T includes the total circuit resistance during the given time interval of consideration. Thus, in the interval t_1 to t_2, here

$$R_T = R_{TF} = R_F + R_1 + R_c$$

where

R_F = forward resistance of the rectifier
R_1 = load resistance
R_c = coil resistance

Here the flux is unchanging at the value ϕ_N, the negative flux swing. Thus, $\frac{d\phi}{dt}=0$ from t_1 to t_2. Then the current becomes

$$i_{1-2}=\frac{Em}{R_{TF}}\cos\omega t \qquad (4)$$

At the time t_2 in the cycle, the proportionality factor between the current and flux has changed so that the differential equation in the interval t_2 to t_3 becomes

$$Em\cos\omega t=\frac{s\phi}{N_1}R_{TF}+\frac{h_0}{N_1}R_{TF}+N_1\frac{d\phi}{dt}\times 10^{-8} \qquad (5)$$

When the time t_3 has been reached in the cycle the flux has reached saturation so that from t_3 to t_4 to t_5, $\phi=\phi_s$ and $\frac{d\phi}{dt}=0$ hence

$$i_1=\frac{Em}{R_{TF}}\cos\omega t \qquad (6)$$

As the cycle of operation proceeds to the interval from t_5 to t_6 the flux has left saturation and is thus changing. The differential equation for this interval becomes

$$E_m\cos\omega t=\frac{s\phi}{N_1}R_{TF}-\frac{h_0}{N_1}R_{TF}+ N_1\frac{d\phi}{dt}\times 10^{-8} \qquad (7)$$

Equation 7 holds until time t_6 where the rectifier resistance changes from the forward resistance R_F to the back resistance, R_{TF} must be changed to R_{TB} in equation 7.

From time t_6 to t_7

$$E_m\cos\omega t=\frac{s\phi}{N_1}R_{TB}-\frac{h_0}{N_1}R_{TB}+ N_1\frac{d\phi}{dt}\times 10^{-8} \qquad (8)$$

At time t_7 the core flux has reached its negative swing and from t_7 to t_8,

$$\phi=\phi_N,\ \frac{d\phi}{dt}=0 \quad \therefore \quad i_1=\frac{E_m}{R_{TB}}\cos\omega t \qquad (9)$$

From these differential equations and instantaneous current relationships the following quantities will be obtained in view of their importance for appropriately describing the action of the circuit:

1. Firing time t_3 as a function of circuit parameters.
2. Time t_5 when the core leaves saturation.
3. Time t_6 when the rectifier leaves forward conduction.
4. Magnitude of negative flux swing.
5. Instantaneous current functions for the complete cycle.
6. Average current for a cycle.

The quantities 1 through 6 can be evaluated from a consideration of the general solutions to equations 5, 7, and 8.

The general solution to equation 5 which represents the flux in period t_2 to t_3 becomes

$$\phi_{2-3}=\frac{C_1}{(C_2^2+\omega^2)}[C_2\cos\omega t+\omega\sin\omega t]- \frac{h_0}{s}+\phi_{c1}e^{-C_2(t-t_2)} \qquad (10)$$

where

$$C_1=\frac{Em}{N_1}\times 10^8$$

$$C_2=\frac{sR_{TF}}{N_1^2}\times 10^8$$

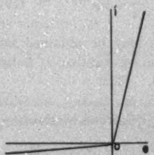


Fig. 2 (left). Rectifier voltage-current characteristic

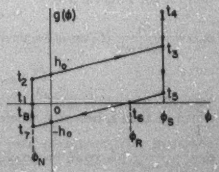


Fig. 3 (right). Magnetization function

ϕ_{c1} = constant of integration

The general solution to equation 7, the equation for interval t_5 to t_6 is

$$\phi_{5-6}=\frac{C_1}{(C_2^2+\omega^2)}\,[C_2\cos\omega t+\omega\sin\omega t]+\frac{h_0}{s}+\phi_{c2}\epsilon^{-C_2(t-t_5)} \quad (11)$$

here ϕ_{c2} is a constant of integration for the interval indicated.

For the interval t_6 to t_7 the solution of the differential equation 8 becomes

$$\phi_{6-7}=\frac{C_1}{(C_4^2+\omega^2)}\,[C_4\cos\omega t+\omega\sin\omega t]+\frac{h_0}{s}+\phi_{c3}\epsilon^{-C_4(t-t_6)} \quad (12)$$

here

$$C_4=\frac{sR_{TB}}{N_1^2}\times 10^8$$

ϕ_{c3} = constant of integration for this interval

As indicated, it is necessary to obtain the t_5 where the core leaves saturation. This time may be obtained by a consideration of the following:

Use of relations 2 and 6 at time t_5 when $\phi=\phi_s$ gives

$$i=\frac{E_m}{R_{TF}}\cos\omega t_5$$

$$\left(\frac{s\phi_s-h_0}{N_1}\right)=\frac{E_m}{R_{TF}}\cos\omega t_5$$

$$\cos\omega t_5=\left(\frac{R_{TF}}{E_m}\right)\left(\frac{s\phi_s-h_0}{N_1}\right) \quad (13)$$

It should be observed from the assumed magnetization function (Fig. 3) that the slope (s) of the sides of the function are given by $s=\frac{h_0}{\phi_R}$. In view of this equation (13) becomes

$$\cos\omega t_5=\left(\frac{R_{TF}}{E_m}\right)\left(\frac{h_0}{N_1}\right)\left(\frac{\phi_s-\phi_R}{\phi_R}\right) \quad (14)$$

This equation is important in that it reveals the dependence of the extinction angle for the core on the core constants and ratio of coercive current to peak load circuit current. This relationship will be discussed more fully in the section devoted to the calculations for various magnetic materials. To obtain the time t_6 it is necessary to utilize the flux function for the interval 5–6.

$$\phi_{5-6}=\frac{C_1}{(C_2^2+\omega^2)}\,[C_2\cos\omega t+\omega\sin\omega t]+\frac{h_0}{s}+\phi_{c2}\,\epsilon^{-C_2(t-t_5)}$$

which at time t_5 becomes, where $\phi=\phi_s$

$$\phi_s=\frac{C_1}{(C_2^2+\omega^2)}\,[C_2\cos\omega t_5+\omega\sin\omega t_5]+\frac{h_0}{s}+\phi_{c_2}$$

and

$$\phi_{c_2}=\phi_s-\frac{h_0}{s}-\frac{C_1}{(C_2^2+\omega^2)}\,[C_2\cos\omega t_5+\omega\sin\omega t_5] \quad (15)$$

The time t_6 may now be obtained implicitly from the relation for ϕ_{5-6} by evaluating ϕ_{5-6} at t_6. Thus, when $t=t_6$ and $\phi_{5-6}=\phi_R$

$$\phi_R=\frac{C_1}{(C_2^2+\omega^2)}\,[C_2\cos\omega t_6+\omega\sin\omega t_6]+\frac{h_0}{s}+\phi_{c_2}\epsilon^{-C_2(t_6-t_5)} \quad (16)$$

Although the time t_6 is not explicitly given by 16 it may be obtained by a substitution method which is indicated in a later section. This time t_6 is an important factor in describing the circuit action because it gives the time when the rectifier current is zero after core extinction. The negative swing of the flux may now be evaluated by a consideration of equation 12. To obtain this quantity, it is first necessary to evaluate the constant ϕ_{c_3} which may be done at t_6, thus

$$\phi=\phi_R=\frac{C_1}{(C_4^2+\omega^2)}[C_4\cos\omega t_6+\omega\sin\omega t_6]+\frac{h_0}{s}+\phi_{c_3} \quad (17)$$

or

$$\phi_{c_3}=\phi_R-\frac{C_1}{(C_4^2+\omega^2)}\,[C_4\cos\omega t_6+\omega\sin\omega t_6]-\frac{h_0}{s}$$

The time t_7 may be obtained by utilizing equations 2, 9, 12, and 17.

$$\frac{Em}{R_{TB}}\cos\omega t_7+\frac{h_0}{N_1}=\frac{s}{N_1}\left[\frac{C_1}{C_4^2+\omega^2}\{C_4\cos\omega t_7+\omega\sin\omega t_7\}+\frac{h_0}{s}+\phi_{c_3}\epsilon^{-C_4(t_7-t_6)}\right] \quad (18)$$

Again equation 18 does not give an explicit relationship for t_7; however, it may be obtained by a substitution method which is explained later.

Since the time t_7 is known, it is now possible to calculate the value of the negative flux swing ϕ_N by a utilization of the flux expression ϕ_{6-7}. When $t=t_7$ the flux ϕ_{6-7} becomes equal to ϕ_N, hence

$$\phi_N=\frac{C_1}{(C_4^2+\omega^2)}\,[C_4\cos\omega t_7+\omega\sin\omega t_7]+\frac{h_0}{s}+\phi_{c_3}\epsilon^{-C_4(t_7-t_6)} \quad (19)$$

The firing time t_3 may be obtained by first calculating the constant ϕ_{c_1} at t_2 by use of equation 10. Then

$$\phi_{2-3}=\phi_N=\frac{C_1}{(C_2^2+\omega^2)}\,[C_2\cos\omega t_2+\omega\sin\omega t_2]-\frac{h_0}{s}+\phi_{c_1}$$

but the value of t_2 must be known. This quantity may be obtained by use of equations 2 and 4. Observe that at t_2

$$\frac{s\phi_N}{N_1}+\frac{h_0}{N_1}=\frac{E_m\cos\omega t_2}{R_{TF}}$$

from which

$$\cos\omega t_2=\left(\frac{R_{TF}}{E_m}\right)\left(\frac{s\phi_N}{N_1}+\frac{h_0}{N_1}\right)=\left(\frac{R_{TF}}{E_m}\right)\left(\frac{h_0}{N_1}\right)\left(\frac{\phi_N}{\phi_R}+1\right) \quad (20)$$

Use of equation 20 will allow ϕ_{c_1} to be calculated from the foregoing.

$$\phi_{c_1}=\phi_N+\frac{h_0}{s}-\frac{C_1}{(C_2^2+\omega^2)}\,[C_2\cos\omega t_2+\omega\sin\omega t_2] \quad (21)$$

The time t_3 may now be obtained by use of equation 10 when $\phi=\phi_s$ at t_3. Thus

$$\phi_s=\frac{C_1}{(C_2^2+\omega^2)}\,[C_2\cos\omega t_3+\omega\sin\omega t_3]-\frac{h_0}{s}+\phi_{c_1}\epsilon^{-C_2(t_3-t_2)} \quad (22)$$

Equation 22 does not explicitly relate the firing point t_3 for the core, however, this time may again be obtained by the substitution method.

To compute the average value of the load current per cycle, it is necessary to compute all of the instantaneous currents which hold in the periods of core and rectifier consideration. The currents which hold in the intervals t_1–t_2, t_3–t_5, and t_7–t_8 have already been calculated and given by equations 4, 6, and 9 respectively. The currents which have not been specified are those in the interval t_5–t_7, and t_2–t_3. The current from t_5 to t_7 must be broken into two parts since the rectifier changes resistance at time t_6. Thus the current from t_5 to t_6 is

$$N_1i_1=s\phi-h_0$$

$$i_1=\frac{s\phi}{N_1}-\frac{h_0}{N_1},\quad \phi=\phi_{5-6}$$

$$i_{5-6}=\frac{s}{N_1}\left[\frac{C_1}{(C_2^2+\omega^2)}\{C_2\cos\omega t+\omega\sin\omega t\}+\frac{h_0}{s}+\phi_{c2}\times\epsilon^{-C_2(t-t_5)}\right]-\frac{h_0}{N_1} \quad (23)$$

The constants c_1, and c_2, and ϕ_{c_2} have already been evaluated. Hence equation 23 provides an explicit expression for the interval (exuations 5 and 6).

In the interval t_6 to t_7 the rectifier resistance changes from R_F to R_B, and the flux function becomes $\phi=\phi_{6-7}$ which gives for the current

$$i_{6-7}=\frac{s}{N_1}\left[\frac{C_1}{(C_4^2+\omega^2)}\,[C_4\cos\omega t+\omega\sin\omega t]+\frac{h_0}{s}+\phi_{c_3}\,\epsilon^{-C_4(t-t_6)}\right]-\frac{h_0}{N_1} \quad (24)$$

The constants c_1, c_4, and ϕ_{c_4} have been obtained, thus equation 24 explicitly defines i_{6-7} the load current which flows in the interval (equations 6 and 7).

The current t_2 to t_3 may be expressed in a similar fashion by ϕ_{2-3}. Hence

$$i_{2-3}=\frac{s}{N_1}\left[\frac{C_1}{(C_2^2+\omega^2)}\,[C_2\cos\omega t+\omega\sin\omega t]+\phi_{c_1}\epsilon^{-C_2(t-t_2)}\right] \quad (25)$$

The constants c_1, c_2, and ϕ_{c_1} have been previously evaluated so that i_{2-3} is explicitly defined.

Discussion of General Relationships

It is observed that the magnitudes of the core constants ϕ_R, ϕ_s, h_0, and s are important factors in determining the boundary angles for the cycle of core operation. Similarly, the magnitudes of rectifier back resistance and forward resistance R_B and R_F are important factors in determining the same boundary angles.

An example of this dependence exists in the angle ωt_5 which represents the angle in the input cycle where the core leaves saturation. By equation 14 it can be seen that the angle depends on two important factors, since

$$\cos\omega t_5=\frac{R_{TF}}{N_1}\,\frac{h_0}{E_m}\left(\frac{\phi_s}{\phi_R}-1\right)=\frac{R_{TF}\times10^8}{\omega N_1\phi_s}\,\frac{h_0}{N_1}\left(\frac{\phi_s}{\phi_R}-1\right)$$

Here it is assumed that unity excitation of the core is employed, i.e., $\frac{E_m\times10^8}{\omega N_1\phi_s}=1$

$$\therefore\ \cos\omega t_5=\frac{R_{TF}}{\left(\frac{\omega N_1^2}{s}\right)\times10^{-8}}\left(1-\frac{\phi_R}{\phi_s}\right) \quad (26)$$

$$a=\frac{R_{TF}}{\omega N_1^2/s}\times10^8 \qquad b=\left(1-\frac{\phi_R}{\phi_s}\right)$$

The factor a is a factor whose general form occurs throughout the development. As may be readily seen, it is of the form of $1/Q$ where Q is the quality of the instantaneous circuit under consideration. If the reactance factor $\frac{\omega N_1^2}{s}\times10^{-8}$ which is equivalent to

$$\omega L_1=\omega N_1^2\mu\left(\frac{4\pi Ac}{l}\right)\times10^{-9}$$

since $\mu=\frac{1}{s}\left(\frac{l}{4\pi A}\right)$ is large in comparison to the total forward resistance of the load circuit, then the angle ωt_6 is very close to $\pi/2$ which is near the end of the voltage cycle even if $\phi_R<<\phi_s$. However, if the factor is small in comparison with the total forward resistance, then the angle of core extinction can become less than $\pi/2$. This condition is often observed in practice, especially where the value of total forward circuit resistance is large in comparison to the unsaturated reactance factor $\frac{\omega N_1^2}{s}\times10^{-8}$.

The angle ωt_6 is implicitly given by equation 16. To clarify the dependence of ωt_6 on the values of parameters, it can be seen that if the unsaturated reactance factor $\frac{\omega N_1^2}{s}\times10^{-8}$ is large in comparison to R_{TF} then $c_2<<\omega$. Thus by equation 16

$$\phi_R=\frac{C_1}{\omega_2}(\omega\sin\omega t_6)+\frac{h_0}{s}+\phi_{c_2}$$

and from equation 15

$$\phi_{c_2}=\phi_s-\frac{h_0}{s}-\frac{C_1}{\omega}\,\omega$$

since

$$\omega t_5\simeq\frac{\pi}{2}$$

Then

$$\sin\omega t_6=\frac{\phi_R}{\phi_s}$$

since

$$C_1=\frac{E_m}{N_1}10^8=\frac{\omega N_1\phi_s}{N_1} \quad (27)$$

If then $\phi_R=\phi_s$ the angle $\omega t_6=\omega t_5$ and further if $\phi_R<\phi_s$ as in the materials without square flux-current loops, the angle ωt_6 lies beyond $\pi/2$. For example, with Hipersil this angle is approximately 118 degrees and for Supermalloy gives a value of 130 degrees.

Another factor which is of great interest in this circuit analysis is the magnitude of the lower tip of the flux swing ϕ_N. Consider equation 19 for ϕ_N in which the factor C_4 enters as an important quantity. The factor C_4 is again a $1/Q$ or inverse quality factor which is now given by the ratio of the back resistance to the unsaturated reactance factor of the core. Thus, $C_4=R_{TB}/N_1^2/s\times10^8$. If now $C_4>>\omega$, which condition would exist if R_{TB} or the rectifier back resistance were very much greater than $\frac{\omega N_1^2}{s}\times10^{-8}$, then by equation 19

$$\phi_N=\frac{C_1}{C_4}\cos\omega t_7+\frac{h_0}{s}$$

However, since the first term is approximately zero due to the high R_{TB} then

$$\phi_N=\frac{h_0}{s}=\phi_R \quad (28)$$

This indicates that the flux cannot swing below the residual flux if the back resistance of the rectifier is high. It is apparent from the general relations, however, that ϕ_N will swing below ϕ_R if the slope factor s becomes very small allowing $R_{TB}\simeq\frac{\omega N_1^2}{s}\times10^{-8}$. This situation is often observed with Supermalloy which has an extraordinarily small value of s. If a large number of load turns N_1 is used with Supermalloy with even the best high back resistance germanium rectifier, the value of ϕ_N is seen to be much less than ϕ_R.

From this relationship it can be appreciated further that the frequency dependence of the quality factor is an important consideration. Especially is this true for the case of small s. At higher supply frequencies the factor $\frac{\omega N_1^2}{s}\times10^{-8}$ becomes large for materials with small s and a usual number of load turns. When this value is compared to values of R_{TB} which are physically realizable it may be appreciated that no advantage is gained in use of the higher slope materials such as Supermalloy at the higher frequencies.

The opposite case that $R_{TB}<<\frac{\omega N_1^2}{s}\times10^{-8}$ yields the relations $C_4<<\omega$, then by equation 19

$$\phi_N=\frac{C_1}{\omega}\sin\omega t_7+h_0/s-\phi_s$$

and by equation 17

$$\phi_{c_3}=-\phi_s\sin\omega t_6$$

also by equation 18

$$\frac{E_1}{R_{TB}}\cos\omega t_7+\frac{h_0}{N_1}=\frac{s}{N_1}\left[\frac{C_1}{\omega}\sin\omega t_7+\frac{h_0}{s}-\phi_s\sin\omega t_6\right]$$

or

$$\cos\omega t_7=\frac{R_{TB}}{\frac{\omega N_1^2}{s}\times10^{-8}}[\sin\omega t_7-\sin\omega t_6]$$

but since

$$\frac{R_{TB}}{\frac{\omega N_1^2}{s}\times10^{-8}}\simeq0$$

then

$$\omega t_7\simeq\frac{3\pi}{2}$$

This gives

$$\phi_N=-\frac{\omega N_1\phi_s}{\omega N_1}+\frac{h_0}{s}-\phi_s \text{ or } \phi_N=\phi_R-2\phi_s \quad (29)$$

In the case of square loop materials this becomes

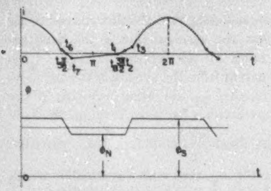


Fig. 4. Flux-current relationship for case of
$$R_{TB}>>\frac{\omega N_1^2}{s}\times 10^{-8}$$

$$\phi_N=-\phi_s \qquad (30)$$

The result (equation 30) indicates that if the back resistance is small in comparison to the unsaturated reactance factor the negative swing of the flux will extend to the lower knee of the flux-current loop. If this is true, then the rectifier no longer offers a difference of resistance to the two directions of applied circuit voltage, hence the average value of the load current is zero and the average value of the flux in the core is zero. Accordingly the major flux-current loop will be traced during 1 cycle of input voltage as indicated by equation (30). Here the limiting cases of large and small back resistance and $\phi_R=\phi_s$ as well as $\phi_R<\phi_s$ have been considered, so as to yield certain information concerning the deterioration of the major loop for the core material as a function of the core and rectifier parameters.

It is of further interest to show that the limiting values of rectifier resistance will yield the proper values of average load current. The two cases to be considered are, for case 1 (see Fig. 4)

$$R_{TB}>>\frac{\omega N_1^2}{s}\times 10^{-8}$$

and for case 2 (see Fig. 5)

$$R_{TB}<<\frac{\omega N_1^2}{s}\times 10^{-8}$$

$$R_{TB}=R_{TF}$$

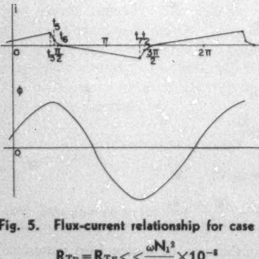


Fig. 5. Flux-current relationship for case of
$$R_{TB}=R_{TF}<<\frac{\omega N_1^2}{s}\times 10^{-8}$$

For case 1 the currents for the complete cycle may be written

$$i_{1-2}=\frac{Em}{R_{TF}}\cos\omega t$$

$$i_{2-3}=\frac{s}{N_1}\left[\frac{C_1}{\omega}\sin\omega t+(2\phi_s+\phi_R)\epsilon^{-C_2(t-t_2)}\right]$$

$$i_{3-5}=\frac{Em}{R_{TF}}\cos\omega t$$

$$i_{5-6}=\frac{h_0}{N_1}\frac{\phi_s}{\phi_R}\sin\omega t$$

$$i_{6-7}=\frac{s}{N_1}\frac{C_1}{C_4}\cos\omega t$$

$$i_{7-8}=\frac{Em}{R_{TB}}\cos\omega t$$

$$t_5\simeq t_6\simeq t_7\simeq\pi/2$$

$$t_1\simeq t_2\simeq 3\pi/2$$

$$\omega t_3=\sin^{-1}\left(\frac{-\phi_R}{\phi_s}\right) \qquad (31)$$

and since

$$I_{1av}=\frac{1}{(t_2-t_1)}\int_{t_1}^{t_2} i_{1-2}dt+\frac{1}{(t_3-t_2)}\times \int_{t_2}^{t_3} i_{2-3}dt/\phi_R=\phi_s,\ t_3=3\pi/2\omega$$

$$I_{1av}=\frac{1}{\pi}\frac{Em}{R_{TF}}-\frac{1}{\pi}\frac{Em}{R_{TB}} \qquad R_{TB}>>R_{TF}$$

$$I_{1av}=\frac{1}{\pi}\frac{Em}{R_{TF}} \qquad (32)$$

This result is completely in agreement with experiment since in an actual measurement it is found that if $\phi_R=\phi_s$ then $I_1=\frac{1}{\pi}\frac{Em}{R_{TF}}$. Such agreement for this case becomes apparent from a consideration of equation 31.

For case 2 the equations for the instantaneous period currents are the same as for case 1 except for the two currents

$$i_{2-3}=\frac{s}{N_1}\frac{C_1}{\omega}(\sin\omega t-\sin\omega t_2\epsilon^{-C_2(t-t_2)})$$

$$i_{6-7}=\frac{s}{N_1}\frac{C_1}{\omega}(\sin\omega t-\sin\omega t_6\epsilon^{-C_2(t-t_6)})$$

which allow the calculation of the average current to proceed on the same basis as for case 1. The average current becomes

$$I_{1av}=-\frac{h_0}{N_1}\frac{\phi_s}{\phi_R}+\frac{h_0}{N_1}\frac{\phi_s}{\phi_R}=0$$

This result of course agrees closely with experiment.

Conclusions

It is important to note that the relative magnitudes of back and front resistance to the unsaturated reactance for the core shall determine for a given core material the quiescent value of the load current (that load current which exists with no external signal applied). The importance of the magnitude of rectifier back-to-front ratio for a core material whose constants are known is shown so that in a practical case the proper core material and load turns may be chosen for a given rectifier, or vice versa, that the material may be utilized in the region of greatest signal sensitivity.

A practical computation of the case of Deltamax material is shown in the following section to illustrate the close agreement with observed results which may be expected from the analysis.

Calculations for Deltamax Core Materials

To illustrate the application of the analysis for a circuit of average practical parameters the following calculations are presented. With reference to Fig. 1 the parameters are

$N_1=2{,}000$ ohms
$R_c=35$ ohms
$R_F=25$ ohms, $R_B=20{,}000$ ohms ... Arnold Engineering core *4178* General Electric type *G-10* rectifier
$R_1=500$ ohms
$R_{TF}=560$ ohms
$R_{TB}=2.05\times 10^4$ ohms
$A_c=0.807$ centimeter^2

The core data for the circuit was obtained by the approximation shown in Fig. 6.

Here

H_0 = coercive ampere-turns per centimeter of core length
h_0 = coercive ampere-turns

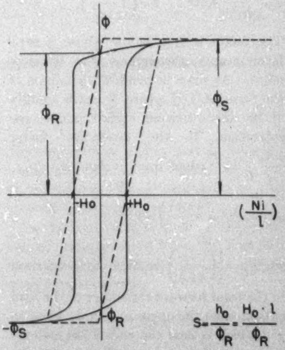

Φ
Φ_R
Φ_S
$-H_0$
$+H_0$
$\left(\frac{Ni}{l}\right)$
$-\Phi_R$
$-\Phi_S$
$S=\frac{h_0}{\Phi_R}=\frac{H_0\cdot l}{\Phi_R}$

Fig. 6. Current-flux loop for Deltamax

Table I. Solution of Equation 33

1	2	3	4	5	6	7	8
θ_6	$\theta_6-90°$	$[1.755\times10^{-4}]$②	1−③	375④	$3.74\cos\theta_6$	$377\sin\theta_6$	⑥+⑦
103°	13	$22.810\times^{-4}$	0.9997	375	−0.842	367	366
95°	5	8.78×10^{-4}	1	375	−0.326	375.5	375.2
95.5°	5.5	9.65×10^{-4}	1	375	−0.35	375.3	374.95

Table II. Solution of Equation 34

1	2	3	4	5	6	7	8	9
θ_7	$\cos\theta_7$	$139\cos\theta_7$	$377\sin\theta_7$	③+④	0.863×10^{-3}⑤	$1.22\times10^{-2}\Delta\theta$	$0.312\epsilon^{-(7)}$	⑥−⑧
240°	−0.5	−69.5	−326	−395.5	−0.341	1.763	0.054	−0.395
245°	−0.423	−58.7	−341.6	−400	−0.345	1.82	0.051	−0.396
246.5	−0.3987	−55.4	−345.7	−401	−0.346	1.85	0.049	−0.395

By interpolation $\theta_7=246.7°$

Table III. Solution of Equation 35

1	2	3	4	5	6
θ_3	$3.53\Delta\theta$	②−281	$100\cos\theta_3$	$9950\sin\theta_3$	④+⑤
360°	316	35	100	0	100
359.6°	314	33	100	−69	31

ϕ_R=residual flux in maxwells
s=ampere-turns per maxwell
A_c=core cross-sectional area in square centimeters
k=core-stacking factor

$$\phi_s=B_skA_c=(14,500)(0.807)(0.85)=9,950 \text{ maxwells}$$

$$\phi_R=B_RkA_c=(14,450)(0.807)(0.85)=9,910 \text{ maxwells}$$

$$s=\frac{H_0l}{\phi_R}=\frac{(0.15)(17.92)}{9,910}=2.71\times10^{-4}\,\frac{\text{ampere-turns}}{\text{maxwells}}$$

h_0=2.69 ampere-turns

$$E_m=\omega N_1\phi_s\times10^{-8}=(377)(2\times10^3)(9,950)\times10^{-8}=75 \text{ volts}$$

Evaluation of all Equation Constants

1. $C_1=\frac{E_m}{N_1}\times10^8=37.5\times10^6$

2. $C_2=\frac{sR_{TF}}{N_1^2}\times10^8=3.79$

3. $C_4=\frac{sR_{TB}}{N_1^2}\times10^8=139$

4. $\frac{C_1}{C_2^2+\omega^2}=26.4$

5. $\frac{C_1}{C_4^2+\omega^2}=23.3$

6. $C_2^1=C_2\frac{\pi}{180}\,\frac{1}{\omega}=1.755\times10^{-4}$

7. $C_4^1=C_4\frac{\pi}{180}\,\frac{1}{\omega}=1.22\times10^{-2}$

The method to be employed in evaluating the various time boundaries and constants of integration will be, first, to determine (θ_5) and then solve successively around the loop, closing the cycle again at (θ_5). This is necessary because only (θ_5) is known explicitly.

8. From equation 13

$$\cos\theta_5=4.05\times10^{-4}$$

Hence

$$\theta_5\approx90°$$

9. From equation 15 ϕ_{c2} is solved at $\theta=\theta_5$

$$\phi_{c_2}=-9,910 \text{ maxwells}=-\phi_R$$

10. From equation 16 θ_6 is solved at $\phi=\phi_R$. Since $C_2^1(\theta_6-\theta_5)$ is very small, the exponential may be expressed by the first two terms of its power series expansion. Numerically, this equation becomes

$$375[1-1.755\times10^{-4}(\theta_6-90°)]=[3.79\cos\theta_6+377\sin\theta_6] \quad (33)$$

This equation must be solved by the method of successive approximations because of its transcendental nature. (See Table I.)

Interpolating, $\theta_6=95.5°$

11. From equation 17 ϕ_{c_3} is evaluated.

$$\phi_{c_3}=-8,430 \text{ maxwells}$$

12. From equation 18 θ_7 is obtained. Numerically

$$\cos\theta_7=0.863\times10^{-3}[139\cos\theta_7 \quad 377\sin\theta_7]-0.312\epsilon^{-1.22\times10^{-2}(\theta_7-\theta_6)} \quad (34)$$

The last term may be neglected if the back resistance is high. (See Table II.)

13. From equation 19 ϕ_N may be determined at $\theta=\theta_7$

$$\phi_N=281 \text{ maxwells}$$

14. From equation 20 θ_2 may be computed at $\phi=\phi_N$

$$\cos\theta_2=1.03\times10^{-2} \text{ or } \theta_2\approx270.6°$$

15. From equation 21 ϕ_{c_1} is derived at $\theta=\theta_2$

$$\phi_{c_1}=\phi_N+\phi_R+\phi_s=20,140 \text{ maxwells}$$

16. From equation 22 θ_3 the firing angle is derived . The last term of this relation may be expanded in a power series. Numerically

$$-281+3.53(\theta_3-\theta_2)=100\cos\theta_3+9,950\sin\theta_3 \quad (35)$$

(See Table III.)

It is evident that the core will be in saturation from θ_3 to θ_5. Hence, the cycle is completed when $\theta=\theta_5$.

Determination of the Instantaneous Currents

1. $90°\leqq\theta\leqq95.5°$

From equation 23

$$i_{5-6}=13.55\times10^{-6}\cos\theta+1.35\times10^{-3}\sin\theta-1.342\times10^{-3}[1-1.755\times10^{-4}(\theta-\theta_5)]$$

θ=90............93 ....95.5
i_{5-6}=0.008........ 0.006......... 0.0017

where θ is in degrees and i is in milliamperes throughout.

2. $95.5°\leqq\theta\leqq246.7°$

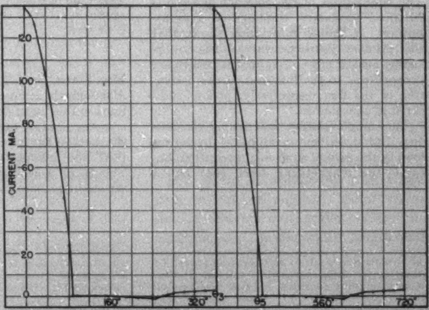


Fig. 7. Load current

From equation 24

$$i_{6-7}=[0.439\cos\theta+1.19\sin\theta-1.042e^{-1.22\times10^{-2}(\theta-\theta_6)}]\times10^{-3}$$

$\theta=$	130	150	170	190	210	246.7
$i_{6-7}=$	−0.056	−0.321	−0.646	−0.968	−1.233	−1.431

3. $246.7^\circ \leqq \theta \leqq 270^\circ$

From equation 9

$$i_{7-8}=3.66\times10^{-3}\cos\theta$$

$\theta=$	250	260	270
$i_{7-8}=$	−1.25	−0.0635	0

4. $270^\circ \leqq 0 \leqq 270^{\circ +}$

Since the current is very close to zero here no attempt was made to determine these currents accurately. Experimentally this assumption was found to be valid.

5. $270^\circ \leqq \theta \leqq 359.6^\circ$

From equation 25

$$i_{8-1}=[0.01355\cos\theta+1.35\sin\theta+2.73(1-1.755\times10^{-4}\Delta\theta)]\times10^{-3}$$

$\theta=$	280	300	320	340	359.6
$i_{8-1}=$	1.40	1.57	1.85	2.247	2.69

6. $359.6^\circ \leqq \theta \leqq 90^\circ$

This is the conducting interval of the cycle. From equation 6

$$i_{1-6}=133.9\times10^{-3}\cos\theta$$

$\theta=$	360	20	40	60	70	80
$i_{1-6}=$	133.9	125.8	102.6	67	45.8	23.2

Shown in Fig. 7 is a graph of the current for two cycles. The experimental curve was found to be in close agreement with this figure.

References

1. SELF-SATURATION IN MAGNETIC AMPLIFIERS, **W. J. Dornhoefer.** *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 835–50.

2. MAGNETIC AMPLIFIERS, **J. R. Walker, D. G. Wilson.** *AFTR 5881*, United States Air Force, Washington, D. C., 1949.

3. THE TRANSDUCTOR AMPLIFIER (book), **U. Krabbe.** *Lindhska Boktryckesiet*, Oribro, Sweden, 1947.

4. DETERMINATION OF STEADY-STATE PERFORMANCE OF SELF-SATURATING MAGNETIC AMPLIFIERS, **Edward J. Smith.** *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1309–17.

5. EVALUATION OF CORE MATERIALS FOR MAGNETIC AMPLIFIERS, **D. C. Dieterly.** *Electrical Manufacturing*, New York, N. Y., Jan. 1953, pp. 68–73.

6. DYNAMIC HYSTERESIS LOOP MEASURING EQUIPMENT, **H. W. Lord.** *AIEE Transactions*, vol. 71, pt. I, Sept. 1952, pp. 269–72.

◆

Discussion

R. M. Wundt (Wright-Patterson Air Force Base, Dayton, Ohio): The choice of the magnetization function is of prime importance in the analysis of magnetic amplifiers. The authors' analysis is based on the ampere-turn function

$$g(\phi)=i_1N_1=s\phi\pm h_0 \qquad (1)$$

This type of function is discontinuous, and represents a box-type hysteresis loop. It is quite suitable to represent core materials with a rectangular hysteresis loop.

Another simple magnetization function can be obtained by considering the exciting current of the core. This is composed of two components, the magnetizing current which is in phase with the flux, and the loss current which is in quadrature with the flux. These two components can be represented by two terms which are respectively proportional to the flux itself, and to the derivative of the flux

$$g(\phi)=c_0\phi+c_1d\phi/dt \qquad (36)$$

The coefficients c_0 and c_1 can be easily determined from the magnetizing current and the loss current of the material, or from the dynamic hysteresis loop taken with sinusoidal current. The same type of magnetization function is obtained by considering the eddy currents in the core material. The laminations of the core act like short-circuited secondary windings of a transformer. A simple calculation shows that the primary ampere-turns of the transformer consist of two components which are proportional to the flux and to the derivative of the flux respectively, just as indicated in equation 36.

This magnetization function is continuous, and describes an elliptical hysteresis loop in the case of sinusoidal excitation, because the magnetizing force $g(\phi)$ has a phase shift with regard to the flux. Hence, hysteresis loops which are more rounded, such as are displayed by Supermalloy, may be represented by the elliptical function of equation 36.

The characteristic differences between the functions of equations 1 and 36 are these:

1. In the function of equation 1 the magnetomotive force undergoes discontinuous changes $\pm h_0$ at certain intervals. This switching of the coercive force h_0 is equivalent to a switching of the control current from $+I_2$ to $-I_2$, where $I_2N_2=h_0$.
2. In the function of equation 36, the magnetomotive force undergoes a continuous change which is proportional to $d\phi/dt$. This function gives the same results as the function of equation 1 when $c_1=0$ and a control current $I_2=h_0/N_2$ is introduced.

An analysis of a single-core 2-mesh circuit (including a control circuit) has been made by using the magnetization function of equation 36. The numerical results are quite similar to those given in this paper. Due to the continuity of equation 36, the times t_2 and t_7, as marked in Fig. 3, do not exist. This brings about a slight simplification of the procedure.

The magnetization function of equation 36 can be further extended to include higher derivatives of ϕ

$$g(\phi)=c_0\phi+c_1d\phi/dt+c_2d^2\phi/dt^2+\ldots \qquad (37)$$

This type of function may be useful if closer approximation of the actual hysteresis loop is desired.

Max Frank, Saul Rabotnick, J. R. Walker: Although the simple analytical method outlined in this paper does not yield new circuit response data, it should be appreciated that the method has been extended to more practical and complex circuit considerations. The first application of this procedure was attempted with the single-core double-mesh amplifier in which the same flux-current function was employed.

The comparison of theoretical and experimental data shows exact agreement; however, it has also been shown that the flux-current function suggested by Dr. Wundt also yields good agreement. Further, the method of analysis suggested in this paper has been extended to the 2-core full-wave doubler amplifier in which some disagreement has been shown between theoretical and experimental data for the case of low rectifier back resistance and square-loop core material. It is thought that the present inaccuracy may be removed by the use of a more realistic representation of the flux-current function, but whether the improvement provided by the addition of a flux-time derivative to the presently assumed function removes this inaccuracy is questioned.

Whether the sinusoidal current excitation which is necessary for the determination of the constants C_0 and C_1 represents the type of core excitation which would occur in actual circuit function is further open to question. It has been our experience that a sinusoidal voltage excitation of the magnetic material will provide data which is more nearly descriptive of the flux-current function occurring in amplifier operation.

In a practical sense, there occurs a possibility regarding the reduction in complexity of an interval analysis of the type contained in this paper for the case of a seemingly continuous core-excitation function of the cut elliptical type, instead of the discontinuous function which was employed. For the single-core 2-mesh amplifier study, it was observed that some reduction in the number of equations necessary for full-cycle representation was possible. However it is felt that such a consideration is not sufficient to warrant the use of this function, which does not extend the accuracy of predicted circuit response.

Digital Computers as an Aid in Electric-Machine Design

R. M. SAUNDERS
MEMBER AIEE

DIGITAL and analogue computers are being recognized gradually as aids in design procedures applied to machines. Recently the Motor Network Calculator (MONECA), an analogue computer for single phase motor design, was described.[1,2] In addition there have been other references describing the use of network and differential analyzers to solve some problems touching on machine design. Thus far there has been very little reference in the literature directly pertaining to the use of digital computers in the machine design function. The purpose of this paper is to attempt to show that under some circumstances the digital computer has a definite place in the design procedure and to attempt to stimulate the use of this valuable tool for design activities.

The Case for the Digital Computer*

The design function most often exercised in electric machine synthesis is the principle of repeated analysis. In practice the designer reviews the specifications for a given machine and then consults either a similar design or a set of curves relating previous experience to the situation at hand. With these data, he applies a certain degree of artistry and makes a series of choices of conductors, slots, winding arrangements, and materials. Having settled on a key group of choices, he analyzes his machine and sees, in the end, if he has met the specifications. If he has met the original conditions, he has a solution. This does not mean that he has the optimum solution. To optimize, he must make repeated analyses until he has sufficient data to plot a control curve or to clearly illustrate that one of the solutions is the optimum. If the first solution does not meet the specifications, he must make repeated analyses until his results converge to the specifications. The number of iterations necessary is a function of the competence of the person doing the work. The more experienced engineer, in general, will converge to a solution with fewer iterations than an inexperienced individual. What the digital computer will do is make these repeated solutions with a great deal of ease and speed, thus releasing the experienced engineer from the drudgery of the computations and leaving him free for more productive work. Furthermore, it will seek out solutions and make repeated analyses at a prodigious rate so that, in effect, "no stone is left unturned."

The basic advantage of the digital computer over hand operations is the speed at which successive iterations can be performed. It is estimated that 25 experienced calculator operators will be about equal to one punched-card calculator. However, the programming or setup time with the punched-card system is sometimes quite long. Thus one must have a considerable volume of work to be done to realize a net saving. To make punched-card calculations attractive, the situation must exist where the final design can be arrived at by repetitive computations of fairly large magnitude. Situations of this type exist where there is a large dollar volume of a given machine or line of machines involved. They also exist in the case of mass-produced items. An example of the first might be a machine for a large generating station where there are several duplicate machines. An example of the second situation is one where a manufacturer plans to build to one design several thousand fractional horsepower motors.

The decision as to whether or not one should build a computer for the specific job is generally made on the basis of the use factor such a machine is apt to have. Veinott[1,2] has described a very fine analogue computer that solves many of the day-to-day problems of small motor design that face his company. Many design situations are such that the volume of work needing a special machine of this type is not sufficient to warrant the investment. In such cases, general purpose digital and analogue computers may also fill the bill. On the other hand, if the use factor on general purpose machines becomes too high, the purchase or building of a special machine might be warranted so that the general purpose machines are made available only for special short-term applications.

The advantages of the digital over the general-purpose analogue computer in this application, which may or may not exist in any given plant, are:

1. Flexibility and economy,
2. Availability,
3. "Naturalness."

The general-purpose digital computer is not designed for any particular computation. Thus a digital computer may at one time be doing accounting computations and in 30 seconds or less be doing machine-design problems. The only shift that needs to be made is the change of a plug board, a tape, or other direction storage equipment.

General-purpose punched-card computers are common installations in businesses and are used extensively for accounting purposes. It has been ascertained* that approximately 25 per cent of all punched card installations have calculating equipment capable of performing the operations described herein. If it may be assumed that a majority of electrical engineers are with large manufacturers, then the percentage of those electrical engineers who have access to punched-card calculators might well be 50 per cent or more. The principles of operation are not difficult to master, and in not more than 10 days concentrated study one should become reasonably competent with the machines.

The digital approach to design is a "natural" because most variables associated with machine synthesis are discrete. Table I shows some of the independent variables present in the design of a synchronous machine. An absolutely discrete variable is one which changes only in definite steps as opposed to continuously variable function. For example, the thickness of steel is graduated by gauge numbers, and this represents a fixed change in thickness for each change in gauge. The number of circuits is another instance; the circuits may be 1, 2, 3, 4, etc., as opposed to 3.28 or any other in-between number. Some variables are discrete only under certain circumstances. The slot width, length, and

---

* By digital computer the author is referring to punched card or more rapid machines; the desk calculator is not considered in this paper.

Paper 54-168, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted October 19, 1953; made available for printing November 27, 1953.

R. M. Saunders is with the University of California, Berkeley, Calif.

The techniques described herein were largely developed while the author served as consultant to the Aeronautical Research Department, Minneapolis Honeywell Regulator Company, Minneapolis, Minn.

* Private communication to the author from J. W. LaForte, Manager, Market Analysis Department, International Business Machines (IBM) Corporation.

Table I. Variables in Synchronous Machine Design

Absolutely Discrete	Discrete Under Certain Circumstances	Continuously Variable
Grade of steel	Outside diameter	Gap diameter
Thickness of steel	Slot width	Axial length
Number of ducts	Slot length	Slot position
Size of ducts	Slot shape	Air gap
Total slots	Pole shape	
Number of circuits		
Pitch of coil		
Conductor size		
Conductor material		
Number of conductors		

Table II. Independent Variables

Identity Number	Cage Description		X_{lm}	R_{rm}	K_p
	Bar Material	End-Ring Material			
01	Copper	Copper	3.50	1.19	0.964
02	Brass	Brass	3.50	5.11	0.964
03	Brass	Copper	3.50	3.45	0.964
04	Half brass Half copper	Copper	3.50	2.28	0.964
05	Copper	Brass	3.50	2.83	0.964
06	Aluminum	Aluminum	3.50	1.90	0.964
07	Aluminum	Copper	3.50	1.59	0.964
08	Aluminum	Brass	3.50	1.89	0.964
09	Half aluminum Half copper	Copper	3.50	1.39	0.964
10	Half brass Half aluminum	Brass	3.50	4.12	0.964

Table III. Operations and Sample Calculations for Ascertaining Performance of 1/2-Horsepower Induction Motor

Step	Operation	Sample Calculation
P 1	S	0.970
2	$(1)^2$	0.941
3	$1-(2)$	0.059
4	K_p^2	0.928
5	$[2-(4)]\,R_{rm}$	1.27
6	$(3)+(5)$	1.33
7	$(3)\times X_{lm}$	0.206
8	$48.2+X_{lm}$	50.0
9	$R_{rm}/(8)^2$	0.0238
10	$(2.00+R_{rm})\times(9)$	0.0760
11	$(7)-(10)$	0.130
12	$(6)^2+(11)^2$	1.78
13	$(8)\times 110$	6.50
14	$110/(8)$	2.20
15	$(14)\times R_{rm}$	2.62
16	$(15)\times(9)$	0.0624
17	$(13)-(16)$	6.44
18	$0.218\times(6)$	0.290
19	$(17)+(18)$	6.73
20	$(15)\times 2$	5.24
21	$(19)^2+(20)^2$	72.8
P 22	$(21)/(12)=I_1^2$	41.0
23	$110\times K_p\times(3)$	6.25
24	$(15)\times K_p$	2.52
25	$(23)^2+(24)^2$	45.4
26	$(25)/(12)$	25.5
27	$(24)^2\times(1)^2$	6.00
28	$(27)/(12)$	3.35
29	$(110\times K_p)^2\times R_{rm}$	13,400
30	$(3)\times(29)$	790
31	$(24)^2\times R_{rm}$	7.58
32	$(30)-(31)$	782
33	$(22)\times 1.0$	41.0
34	$(26)\times R_{rm}$	30.4
35	$(28)\times R_{rm}$	4.00
36	$(32)\times(2)/(12)$	413
P 37	$(33)+(34)+(35)+(36)+24.0$	512
P 38	$(36)-24.0-17.0$	372
39	$(1)\times 1800$	1,746
P 40	$112.6\times(38)/(39)$	24.0
41	$(38)/(37)$	0.726
P 42	$(37)^2/110^2\times(22)$	0.528
P 43	$(41)^2\times(42)$	0.279
P 44	$(38)/0.5\times 0.746$	0.997

Step 1. Per-unit speed
Step 22. (Input current)²
Step 37. Power input
Step 38. Power output
Step 41. Torque
Step 42. (Power factor)²
Step 43. (Apparent efficiency)²
Step 44. Per-unit load

shape may be continuously variable if the slot die has not yet been fabricated. In many instances, the designer must use existing slot dies and hence is restricted to certain discrete changes. The same may be said of the pole shape. The outside diameter may be continuously variable within certain limits, but it is more common to fix the outside diameter to the standard frame sizes that a given concern uses. The air-gap diameter, and the air-gap and the axial length are the only really continuous variables present. Hence, changes in design variables are made in steps and lend themselves to digital computation.

General Procedure

Having decided that a digital computation is called for, the steps to be followed are as follows:

1. Ascertain the factors to be varied and the limits over which they are to be varied.
2. Establish relationships connecting the factors together.
3. Decide upon figures of merit.
4. Make a table of operations and run through a sample calculation by hand. (Normalization may be helpful.)
5. Code the solution for the machine.
6. Run the cards and print results.
7. Interpret results.

In long computations it is sometimes desirable to decide on intermediate figures of merit so as to cull out undesirable designs early in the computation. While this procedure saves time, there is the danger that a good design may be rejected by faulty figures of merit.

The author has used punched card techniques for the solution of problems involving:

1. Field distribution;
2. Optimum design for a synchro;
3. Optimum cage design for an induction motor.

The field distribution problems were of two types: to find a winding configuration that would produce a given field distribution, and to find the coefficients of a Fourier series representing the permeance of the air gap of a machine whose slot openings contributed sizable harmonics. The synchro design problem was one in which an optimum design was being searched for in terms of fixed large volume production. This was a large-scale problem during which the complete design was set up for punched-card analysis and 37,000 possible designs were considered through use of the digital computer. The cage-design problem is one which is used as a classroom illustration of what can be done with digital techniques when applied to the classic circuit-analysis problem. The problem chosen in this paper is that of calculating the speed-torque curves for an induction motor as a function of cage material, bar size, and bar spacing. It appears in a textbook, Kuhlmann[3]; thus being readily available to all who wish to try their hand at a digital-computer problem. The complete treatment of the problem is contained in the Appendix.

Conclusions

The author does not consider himself an expert in digital computing techniques and feels sure that the procedures he has used may be improved upon by people better informed on digital computers than he. However, he is convinced that the digital computer in the hands of an ordinary design engineer is a powerful tool in performing calculations whose magnitude makes them otherwise impractical.

The author also feels that these new computing tools are going to assist materially in making our machines more economical to build and in making them perform better for the same weight and cost. These are ever the goals of design engineers: digital computers will make the achievement of the objective more rapid.

Appendix

In the following are the steps needed to solve an optimum cage design problem for a 1/2-horsepower single-phase induction motor. This particular problem is used for class demonstration purposes for illustrating digital-computer techniques. As it stands, the magnitude of the problem is not worth while setting up on a punched-card calculator; it could be solved in less time by other techniques. It should be pointed out, however, that once the solution is coded and the plug boards set up, they may be left that way and used over and over again. Thus, by tying up one or two boards continuously, one may always have at his disposal rapid methods of finding performance characteristics of any motor he may design in the future. The value of such a technique is not to be overlooked, for the board represents a small investment and the machine may be used for other purposes while the motor-calculation boards are idle. This is in direct contrast to the construction of a special computer that may be used for this one purpose only.

As was mentioned the motor chosen for this example is one whose whole design is available.[3] The problem whose solution is illustrated here might be to find which of the several cages proposed in Table II will yield the maximum torque at any speed, that cage which will yield maximum speed at rated torque, or to generate data for performance curves. In Kuhlmann,[3] the method of performance calculation is given in tabular form; several parameters which do not vary with speed are given at the top of the table and then computations involving these parameters with the chosen speed are given at the lower portion of the page. By rearranging the table somewhat so that the parameters at the top of the page could be included in the table, the number of input quantities can be reduced to three independent variables. Table II shows the independent variables, X_{1m}, R_{rm}, and K_p, which are constant for all speeds, but change with bar size or composition. These are entered, together with a code or identity number, on cards. Then the speeds are chosen at which the computations are to be made. The next step is to prepare the operation and sample-calculation sequence, Table III.

In Table III, the same sample calculation is made as Kuhlmann[3] makes, with the exception of the input current, the power factor, and the apparent efficiency. These changes are necessitated by the difficulty in extracting square roots by the calculator. Thus the quantities mentioned are left in squared form. These may be extracted rapidly by slide rule or left in this form, as the engineer wishes. Quantities which are to be punched are marked with a P to the left. Thus, the quantities punched are:

1. Per-unit speed, step 1.
2. (Input current),^{2} step 22.
3. Power input, step 37.
4. Power output, step 38.
5. Torque, step 41.
6. (Power factor),^{2} step 42.
7. (Apparent efficiency),^{2} step 43.
8. Per-unit load, step 44.

After the cards are run, and the results punched and printed at the top of the card, they may be sorted for maxima or minima, or otherwise shuffled for various considerations. Data may be read directly from the card, or they may be printed on summary sheets.

References

1. MONECA—A NEW NETWORK CALCULATOR FOR MOTOR PERFORMANCE CALCULATIONS, **Cyril G. Veinott**. *AIEE Transactions*, vol. 71, pt. III, 1952, pp. 231–39.

2. MONECA—A NEW NETWORK CALCULATOR FOR MOTOR PERFORMANCE CALCULATIONS, **C. G. Veinott**. *Electrical Engineering*, vol. 71, Sept. 1952, pp. 795–801.

3. DESIGN OF ELECTRICAL APPARATUS (book), **J. Kuhlmann**. John Wiley and Sons, New York, N. Y., third edition, 1950, p. 389.

◆

Discussion

Robert L. Fillmore (University of Minnesota, Minneapolis, Minn.): Strengthening the point made by the author that digital computers are a great aid in everyday design problems, a digital computer was used recently to aid in determining the equivalent circuit for the drag-cup induction motor, utilizing test data from three experimental motors each having four or five rotors. By selecting a range of likely values for each parameter of the equivalent circuit, and using three or four particular values to cover this range, a family of 2,800 circuits was calculated for each stator and rotor combination. The best equivalent circuit for each motor was determined by the computer by comparing the results of the calculation with the test results. The machine then tabulated the calculations in order with the best comparison at the top of the list.

From this information it was possible to determine formulas to give the correct parameter for each motor and to set up as design procedure to calculate performance of drag-cup motors. This problem could not be solved in a reasonable time without digital computers. The fact that a computer can compare results and make decisions on the basis of some figure of merit is not brought out clearly in the paper, but is, nevertheless, an added advantage in using the digital computer.

C. G. Veinott (Reliance Electric and Engineering Company, Cleveland, Ohio): Professor Saunders makes reference to a paper I presented 2 years ago.[1] This computer was designed for a specific purpose to which it was well suited and for which it is still being used. Since then, I have gained some actual experience with digital calculators, enough to make some comparisons. The calculator I have been using is an IBM Card-Programmed Calculator (CPC). A comparison properly begins by comparing the relative areas of usefulness of a network-type computer with a digital calculator.

First, the network calculator can be used only for those problems that can be reduced to an equivalent electrical network. In motor design, this means evaluation of the performance from the lumped-circuit constants, but does not include calculation of the constants themselves from the geometry of the machine. Single-phase motors require a great deal of this type of calculation, particularly permanent split capacitor motors. Hence, the assistance a network calculator can afford in single-phase motor design is important. In polyphase motor design, comparatively little use is made of the constants—much more work is involved in computing the constants. For such a case, the network computer is of relatively less use, because the relative area of its usefulness is necessarily limited. A network computer can be most helpful in design of double-cage polyphase motors.

Digital computers have a virtually unlimited area of usefulness, for they can do all that the network calculator can plus a great deal more. In the design of d-c motors, for example, most of the calculations are of a nature that cannot be performed on a network computer, which is therefore worthless for such an application. I had a polyphase motor design procedure programmed on a CPC and a number of calculations made. The input data was the geometry of the punching and the winding specification. The machine computed such data as:

1. Copper weight.
2. Resistances, stator and rotor.
3. Total flux.
4. Areas and densities of the various portions of the magnetic circuits.
5. Exciting current.
6. Reactances.
7. Deep-bar effects.
8. Locked-rotor current and torque.
9. Breakdown torque.
10. Full-load efficiency and power factor.
11. Current density.

and many other factors of importance in evaluating the worth of the design.

From the educational point of view, the network calculator is probably more useful, because the engineer can set up the network directly and can see better what is happening; it is possible to juggle one constant at a time and instantly watch the effect.

Cost is a most important consideration. The IBM CPC rents for approximately 30,000 dollars per year. MONECA, even including development costs, for it was the first of its kind built, costs only a fraction of the annual rental charge of the CPC. Operating and maintenance costs of MONECA have been negligible in the first 2 years of operation. Consequently, an idle MONECA does not cause the concern to a budget-conscious manager that an idle CPC would. In cost comparisons, other types of computers need to be considered.

From his remarks concerning the availability of computers such as he was using, I would infer that probably Professor Saunders was not using a CPC, but an IBM *602A* or *604*. The latter two machines are quite generally available in punched-card installations, because they serve the needs of most accounting departments and are less costly. The CPC is not so generally available. Until recently, digital computers of the type needed cost from 250,000 to 500,000 dollars or more. Recently some magnetic-drum digital computers have been developed which sell outright in the 50,000- to 100,000-dollar range. Some of these appear to have important advantages over the CPC for machine design work of the nature I have been discussing.

For a problem even so simple as straight polyphase induction-motor design, we found 602 and 604 machines unsuitable. Attempts were made to program the calculation schedule, but too many board changes were required to make the use of these lower-cost machines suitable for our purpose. The CPC did the job very satisfactorily.

My own conclusions to date are that the possibilities of the giant-brain type of digital computer for machine design have not been scratched yet; that the need for them will grow with time. Principal difficulties in the way of their greater use now are:

1. High cost of an installation.
2. Excessive size of the necessary apparatus.
3. Human inertia to a change.

These computers could relieve the design engineer of an enormous amount of pick-and-shovel calculations that he now makes, or skips and guesses the result. Transistors may hold the key to smaller physical size; the costs may come down later.

REFERENCE

1. See references 1 and 2 of the paper.

R. M. Saunders: Mr. Fillmore adds another area where the punched-card digital computer has contributed greatly to the ease in solving an otherwise difficult problem. The one he has dealt with is a vexsome problem which designers of servo equipment face daily. By finding the true equivalent circuit from test data, he may modify design equations to yield equivalent circuits close to those from test procedures. Although this problem is more of a research than design nature, it is of interest since it does illustrate the great assistance a digital computer can be to those working in the rotating-machinery field.

Mr. Fillmore also points out the importance of choosing a figure of merit to which the computer can compare any solution it makes. As he says, this choice is a very important one and one which will assist greatly in arriving at the best solution.

Mr. Veinott mentions the difficulty in justifying the cost of punched card operations at the CPC level. While the CPC does rent for approximately 30,000 dollars per year, it will solve the same problem as a smaller installation in much less time. Thus, on an hourly charge basis, the CPC might well be the least expensive. It should also be added that one does not install a CPC without having sufficient work to keep its load factor high. If the load factor is high, then it will undoubtedly be solving problems in other areas as well as motor-design problems. A special computer, either digital or analogue, is idle when not solving the specific problem for which it was designed. Hence, I feel that Mr. Veinott's cost analysis is not entirely fair. His point is well taken, however; the justification must lie in the economic realm.

Calculation of Life Characteristics of Insulation

L. C. WHITMAN
MEMBER AIEE

PAUL DOIGAN
AFFILIATE AIEE

Synopsis: Based on chemical reaction rate theory and an extension of the application of the Arrhenius equation, a method is presented for calculating insulation degradation curves. Using the initial dielectric strength (DS) and time-temperature insulation characteristic to deteriorate to some given percentage level, curves showing the relationship between DS and insulation aging at a constant temperature may be computed. Also degradation curves to other given percentage levels of initial DS may be derived. Thus data may be converted into forms more valuable to the design engineer.

Criterion of Insulation Life

VARIOUS criteria of insulation life have been proposed, each having its special advantages and application. In liquid-filled apparatus mechanical criteria of insulation, such as tensile strength, folding, tearing, etc. have been successfully used. A DS test on liquid-filled apparatus is not a good criterion since the liquid impregnant will replace the degraded solid portions and maintain the DS if undisturbed by mechanical shocks such as short-circuit forces. Dry-type insulations, on the other hand, may maintain their mechanical strength when the impregnant, and often the complete structure, has lost a major portion of its DS.

The DS of dry-type insulation however, may be successfully used on many composite insulations as a measure of degradation and hence of insulation life. The insulation life data used in this paper are based on the time required for the hot DS of the insulation to decrease to some percentage of its hot initial value. Data for both 50 per cent and 66 per cent of the hot initial value are shown.

Test Sample

Insulating materials investigated are 1-mil polyethylene terephthalate (Mylar) and 15-mil asbestos, silicone-impregnated.

The insulating material being investigated is wound on a 2-inch-diameter steel mandrel having an internal electric heating element. Test electrodes are applied over the insulation. Thermocouples on the mandrel and on the electrodes are used to check continuously the aging temperature at each test area. Sixty-cycle dielectric tests in air are made between the electrodes and the mandrel at regular time intervals at 50 per cent or 66 per cent (depending on criterion being used) of the initial hot DS strength. All tests are made at the hot temperature. The arrangement of the test sample is schematically shown in Fig. 1. Further details of this sample and temperature determination are given in a previous technical article.[1]

Theoretical Basis and Equation Derivation

Insulation aging data are frequently presented by plotting the reciprocal of the absolute aging temperature as abscissa and the logarithm of the aging time to failure as ordinate. This gives concise and useful information but often the design engineer requires this information in other forms, for example, aging at a constant temperature or perhaps at a different level of degradation. To show that data can be logically and correctly converted to other forms, some theoretical considerations and derivations are necessary.

A quantitative equation relating the influence of temperature and a specific reaction chemical rate is given by the Arrhenius equation[2]

$$k = A\epsilon^{-E/RT} \qquad (1)$$

where

k = specific reaction rate
E (the activation energy) = constant for a particular reaction
R = universal gas constant
T = absolute temperature and
A = frequency factor of molecular encounter.

Following the reasoning of Dakin,[3] if we can assume that the physical property which we are measuring, namely the DS, is a function of time and that it is proportional to a constituent that is a measure of insulation deterioration in a manner analogous to the concentration of a reacting substance and the chemical reaction rate

$$-\frac{dc}{dt} = kc^n \qquad (2)$$

then we have a basis for discussion.

For a first-order chemical reaction, $n=1$, and integration of equation 2 gives the following

$$\log_e c = -kt + \text{constant} \qquad (3)$$

It is not known how many insulation deterioration processes follow simple first-order reaction mechanisms, but nevertheless a function of concentration can be found to satisfy the integration of the general reaction rate equation.

$$f(DS) = -kt + \text{constant} \qquad (4)$$

If the insulation deterioration measured by the DS follows a first-order reaction, then $f(DS) = \log_e DS$ and a plot of $\log_e DS$ versus time will give a straight line.

$$\log_e DS = -kt + \text{constant} \qquad (5)$$

Once the slope of such a plot is established it is a relatively simple matter to extrapolate constant-temperature short-time test data to longer periods of time.

From equation 1, one can see that the specific reaction rate k is temperature-dependent and thus our measured physical property is shown to be temperature-dependent by equating equations 1 and 5. Then

$$\log_e DS = -At\epsilon^{-E/RT} + C \qquad (6)$$

Taking the logarithm of equation 6 we have

$$\log_e t = \log_e\left[\frac{C - \log_e DS}{A}\right] + \left(\frac{E}{R}\right)\frac{1}{T} \qquad (7)$$

As previously stated, we have selected 50 per cent of the initial hot DS as a reasonable criterion of life. Then $\log_e$ (50-per-cent DS) will be a constant for any given temperature and the time to reach this breakdown value will be the measure of insulation life. A plot of the logarithm of the time to reach a specified percentage of DS value versus the reciprocal of the absolute temperature will generally be a straight line. From such a plot, relatively short-time high-temperature data can be extrapolated to long-time operating temperature limits.

It must be recognized that the data will plot as a straight line only if a single reaction is occurring or if several reactions are occurring simultaneously with similar temperature coefficients. This situation evidently exists for Mylar film

---

Paper 54-72, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted October 15, 1953; made available for printing November 25, 1953.

L. C. WHITMAN and PAUL DOIGAN are with the General Electric Company, Pittsfield, Mass.

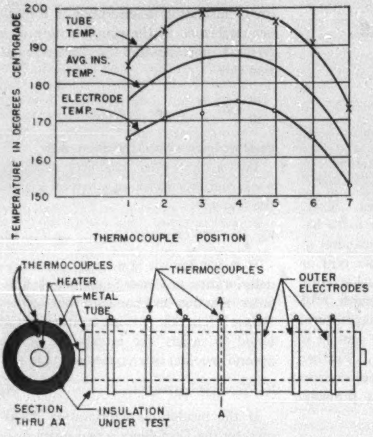


Fig. 1 (above). Test sample and typical temperature gradients

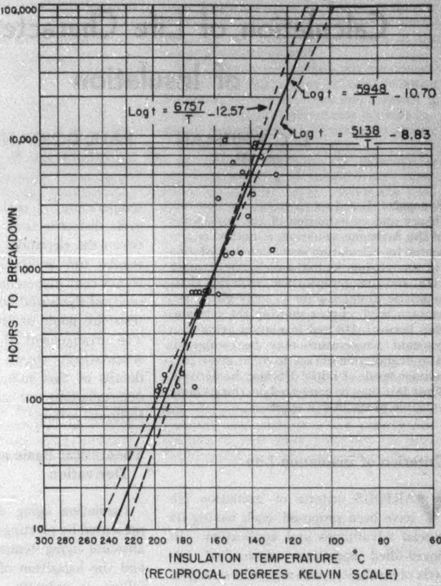


Fig. 2 (right). Tested life characteristics of polyethylene terephthalate film in air, 50-per-cent DS criterion

and silicone-treated asbestos since our data conform to this straight-line relationship.

Subsequent discussions depend on data obtained from the plot of the reciprocal of the absolute temperature versus the logarithm of time to breakdown ($1/K$ versus $\log_{10} t$). Thus, it is extremely important that the straight-line graph from these data be the best obtainable. Any errors made in placing this curve will be magnified in subsequent operations. We have drawn the best obtainable curve by use of the method of least squares in which we calculate both the intercept and slope, using all of the plotted data.

The general form of the equation for a straight line is $y=a+bx$, or in terms of our data

$$\log_{10} t = a + b\frac{1}{T} \tag{8}$$

The constants a and b represent the intercept and slope of the straight line respectively.

By the method of least squares,[4] the equations for the calculation of the intercept and slope are

$$a=\frac{\Sigma y\Sigma x^2-\Sigma xy\Sigma x}{n\Sigma x^2-(\Sigma x)^2} \tag{9}$$

$$b=\frac{\Sigma y\Sigma x-n\Sigma xy}{(\Sigma x)^2-n\Sigma x^2} \tag{10}$$

In using this form we have assumed that the x values (temperature) are free from error and the calculations made to minimize the deviations of y, log time.[4] As in most cases, interchange of co-ordinates will give another straight line. With interchange of co-ordinates equation 8 becomes

$$\log_{10} t=\frac{\frac{1}{T}-a}{b} \tag{8A}$$

These calculations are shown in Appendix I and Table I, and curves as shown in Fig. 2 are the result.

Application of Equations to Polyethylene Terephthalate Insulation

The integration constant in equation 6 can be evaluated by substituting values of $\log_e DS$ at time $t=0$. For the material under consideration, polyethylene terephthalate (Mylar) film, the average initial DS is 3,300 volts per mil at room temperature. Log_e 3,300 (and consequently the integration constant) is equal to 8.10.

To determine the values of A and E/R for use in equation 6 it should be noted that equation 7 is in the same form as equation 8, hence

$$a=\log_{10}\left[\frac{C-\log_e DS}{A}\right] \tag{9A}$$

and

$$b=\frac{E}{2.3R} \tag{10A}$$

Equating equations 9 and 9(A) (using average value of a as calculated in Appendix I) will allow the solving for A

Table I. Polyethylene Terephthalate Film Life Test Data

Degrees Centigrade	Hours = t
127	5,171
129	1,393
135	6,999
138	8,419
138	8,371
139	3,561
142	2,472
145	5,350
147	1,300
150	6,278
152	1,300
155	9,360
155	1,300
156	1,372
158	3,472
159	637
166	637
167	656
171	637
172	637
173	637
175	123
183	153
186	123
194	116
194	153
198	123
198	116

Table II. Silicone-Asbestos Life Test Data

Degrees Centigrade	Hours = t
230	2,675
244	950
252	3,575
256	3,100
254	2,675
245	4,425
226	4,425
260	3,675
270	1,050
280	700
285	700
280	700
269	2,400
252	2,700
310	275
330	275
346	275
352	275
352	275
346	175
328	420
377	3
388	3
395	3
398	3
390	3
375	3
350	15
300	186
315	62
325	27
328	27
320	27
312	86
302	138
355	27
375	4
392	4
397	4
380	4
365	4
345	86

giving a value of 3.47×10^{10}. Note that the term $C-$ $\log_e$ DS is equal to the natural logarithm of the ratio of initial DS (100 per cent) to per-cent criterion of *DS* used. Thus for 50-per-cent DS criterion this term will be $\log_e 2=0.693$. Similarly, equating equations 10 and 10(A) (using average value of b as calculated in Appendix I) will determine E/R which for this case is $1.37=10^4$.

Equation 6 then becomes for polyethylene terephthalate

$$\log_e DS=-3.47\times10^{10}\,t\epsilon^{-\frac{1.37\times10^4}{T}}+8.10 \qquad (11)$$

This equation should now represent a family of curves which shows the relationship between DS and aging at a constant temperature for an insulation material whose electric insulation deterioration follows a first-order rate reaction and whose life characteristics are like those for polyethylene terephthalate (Mylar) film.

Using equation 11, the curves for DS versus time can be plotted for any operating temperature. This has been done for three temperatures, 95, 110, and 130 degrees centigrade (C) and is shown plotted in Fig. 3. The initial hot DS for these values of temperature is essentially constant.

It is possible to ascertain the life characteristics of the insulation at any percentage of the initial DS. If the 50-per-cent criterion is not the best one to use, then the data may be interpreted in terms of a 25-per-cent or 75-per-cent criterion or any other percentage desirable. To illustrate this, the dotted curves in Fig. 4 were computed from equation 13 and plotted. It is interesting to point out that the true value of the initial DS is not important in calculating these curves (but necessary to obtain Fig. 3) since equation 11 contains an integration constant which is the logarithm of the initial DS. Therefore, there will always be a ratio term equal to the ratio of the percentage criteria (the difference of two logarithms is the ratio of the numbers).

The equation giving the relation between per-cent DS, temperature, and hours to breakdown may be derived from equation 7 which becomes

$$\log_e t=\log_e\left[\log_e\left(\frac{100}{\text{per-cent DS}}\right)\right]-\log_e A+\frac{E}{RT} \qquad (12)$$

For polyethylene terephthate, this becomes

$$\log_e t=\log_e\left[\log_e\left(\frac{100}{\text{per-cent DS}}\right)\right]-24.270+\frac{13700}{T} \qquad (13)$$

Verification of the correctness of this analysis and the conversion of data are given by an experimental check made on polyethylene terephthalate (Mylar) insulation using a 66-per-cent criterion. The test points shown in Fig. 4 are for 66-per-cent criterion, and these rather limited data show good agreement with the computed 66.6-per-cent curve.

Application of Equations to Silicone Asbestos Insulation

A similar analysis of data for silicone impregnated asbestos insulation is given in Appendix II. Original data are shown plotted in Fig. 5 with the best-fitting curve.

From the initial DS of 160 volts per mil for the silicone asbestos at room temperature, the integration constant for this material is 5.07 ($\log_e 160$).

From equations 9(A) and 10(A) (using average values of a and b as calculated in Appendix II) the values of $E/R=1.687\times10^4$ and $A=1.64\times10^{10}$ are determined. Equation 6 then becomes for silicone asbestos

$$\log_e DS=-1.64\times10^{10}\,t\epsilon^{-\frac{1.687\times10^4}{T}}+5.07 \qquad (14)$$

This equation should now represent a family of curves which shows the relationship between DS and aging at a constant temperature for an insulation material whose electrical insulation deterioration follows a first-order rate reaction and whose life characteristics are like those of silicone asbestos insulation.

Using equation 14 the curves for DS versus time can be plotted for any operating temperature. This has been done for four temperatures, 180, 200, 225, and 250 centigrade and is shown in Fig. 6. The initial hot dielectric strength for these values of temperature is approximately constant.

The equation giving the relation between per-cent DS, temperature, and

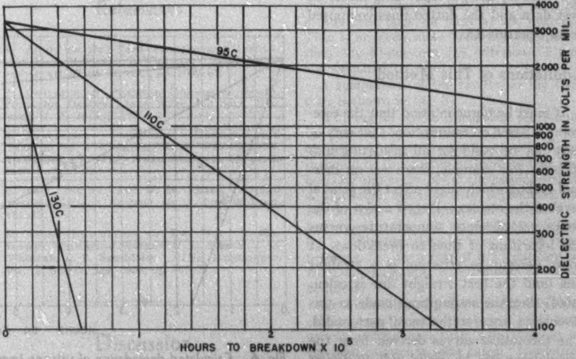


Fig. 3. Calculated degradation of polyethylene terephthalate film in air at constant temperatures, 50-per-cent DS criterion

$$\log_e DS=-3.47\times10^{10}\,t\epsilon^{-\frac{1.37\times10^4}{T}}+8.10$$

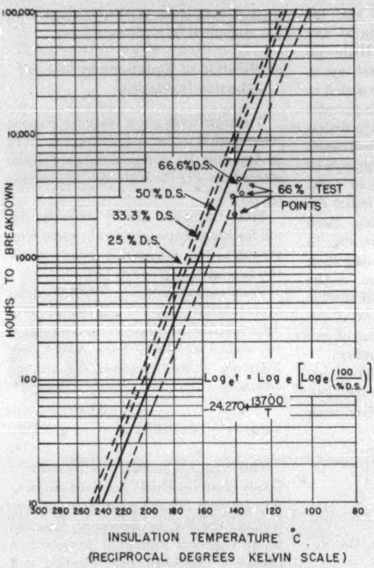


Fig. 4. Calculated and tested life characteristics of polyethylene terephthalate film in air at various DS criteria

hours to breakdown for the insulation by inserting the proper constants in equation 12 becomes

$$\log_e t = \log_e\left[\log_e\left(\frac{100}{\text{per-cent DS}}\right)\right] - 23.520 + \frac{16870}{T} \quad (15)$$

Fig. 7 shows the relation for different DS criteria, the full line being based on test data and the dotted lines computed from equation 15.

Limitations of This Method

It must be borne in mind that the foregoing is based on assumptions that may or may not be valid for all insulating materials. For example, there was no thermal cycling which might affect the general level of the results. But, if a plot of the reciprocal absolute temperature versus the logarithm of time to breakdown at some percentage criterion is a straight line (and the best straight line is calculated) then the assumptions made in this discussion are, for the most part, valid. The theoretical curves derived from the calculations will be useful aids to design engineers in selecting insulating materials.

In selecting a criterion of useful life, one should consider more than just the theoretical aspects. Selecting a 90-per-cent criterion will give us the desired data in less time than it would when we use a 50-per-cent criterion but the accuracy of the data may not be as good. Percentagewise errors will be magnified. On the other hand, data obtained using a 10-per-cent criterion may be more accurate but also would require a much longer testing period. There will be a lower limit to the per-cent DS criteria that can be used if volts per mil of the selected breakdown criteria become less than the electrical strength of the air film of thickness equal to thickness of insulation under test. In this case no breakdowns would occur regardless of time

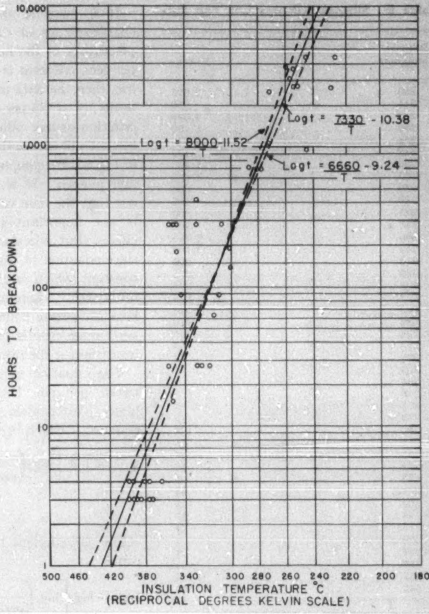


Fig. 5. Tested life characteristics of silicone-impregnated asbestos insulation in air, 50-per-cent DS criterion

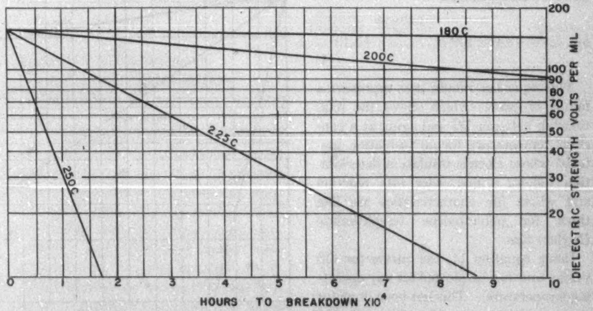


Fig. 6. Calculated degradation of silicone-impregnated asbestos in air at constant temperatures, 50-per-cent DS criterion

$$\log_e DS = -1.64\times10^{10} t e^{-\frac{1.687\times10^4}{T}} + 5.07$$

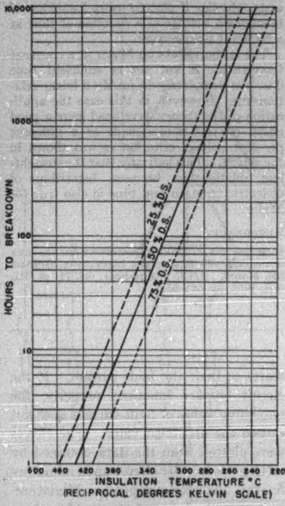


Fig. 7. Calculated and tested life characteristics of silicone-impregnated asbestos in air at various DS criteria

$$\log_e t=\log_e\left[\log_e\left(\frac{100}{\text{per-cent DS}}\right)\right]-23.520+\frac{16870}{T}$$

since air would flow into the regions of deteriorated insulation and maintain this minimum strength. The choice of the criterion must be made with these factors in mind in addition to the properties of the specific insulation.

Conclusions

1. A method is available for converting time-temperature insulation characteristic curves to insulation aging characteristics at a constant temperature.

2. Insulation aging curves at one level of degradation may be extended to other levels if a first-order reaction exists.

3. Extended aging data on polyethylene terephthalate (Mylar) and silicone asbestos confirm more limited data previously published.[5]

4. Limited tests on polyethylene terephthalate (Mylar) at a different percentage criterion confirm the correctness of this method for this type of material.

Appendix I. Analysis of Data for Polyethylene Terephthalate

Computing values of a and b from equations 9 and 10, noting that $x=1/C+273$ and $y=\log_{10} t$, and substituting in equation 8 we obtain

$$\log_{10} t=-8.83+\frac{5138}{T}$$

When the calculations are made interchanging the x and y co-ordinates in 9 and 10 and substituting in equation 8(A) this becomes

$$\log_{10} t=-12.57+\frac{6757}{T}$$

which gives approximately the same straight line. We have taken the average of these two curves as the best straight line for the data as shown in Fig. 2, and the equation becomes

$$\log_{10} t=-10.70+\frac{5948}{T}$$

Appendix II. Analysis of Data for Silicone Asbestos Insulation

Computing the values of a and b from equations 9 and 10 noting that $x=1/C+273$ and $y=\log_{10} t$ and substituting in equation 8 we obtain

$$\log_{10} t=-9.24+\frac{6{,}660}{T}$$

When the calculations are made interchanging the x and y co-ordinates in equations 9 and 10 and substituting in equation 8(A) this becomes

$$\log_{10} t=-11.52+\frac{8{,}000}{T}$$

which gives approximately the same straight line. We have taken the average of these curves as the best straight line for the data as shown in Fig. 5, and the equation becomes

$$\log_{10} t=-10.38+\frac{7{,}330}{T}$$

References

1. AGING OF DRY-TYPE TRANSFORMER INSULATION, **H. C. Stewart, L. C. Whitman.** *AIEE Transactions*, vol. 67, pt. II, 1948, pp. 1600–07.

2. OUTLINES OF PHYSICAL CHEMISTRY (book), **Farrington Daniels.** John Wiley and Sons, New York, N. Y., 1948, p. 377.

3. ELECTRICAL INSULATION DETERIORATION TREATED AS A CHEMICAL RATE PHENOMENON, **T. W. Dakin.** *AIEE Transactions*, vol. 67, pt. I, 1948, pp. 113–22.

4. ONE LINE OR TWO? **W. M. Jessop.** *Applied Statistics*, London, England, vol. 1, June 1952, pp. 131–37.

5. AGING EVALUATION OF DRY TYPE TRANSFORMER INSULATING SYSTEMS, **H. C. Stewart, L. C. Whitman, A. L. Scheideler.** *AIEE Transactions*, vol. 72, pt. III, Apr. 1953, pp. 267–77.

◆

Discussion

T. W. Dakin (Westinghouse Research Laboratories, East Pittsburgh, Pa.): The life test data on silicone impregnated asbestos and Mylar insulation are interesting, as is the scheme for calculating the dielectric strength time relation for any given temperature from data on the variation of the time to reach a certain fraction of the initial dielectric strength as a function of temperature. I feel that we must be careful in using arbitrary fractional amounts of the initial DS as criteria of insulation life. Such life values can only apply to situations where it is the function of the insulation to have such DS levels. Where the insulation needs to have only much less DS or perhaps only serves as a spacer, such life data, using a 50-per-cent DS end point are apt to be unrealistic and pessimistic.

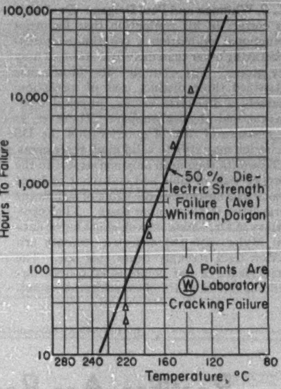


Fig. 8. Life tests on Mylar film

It is interesting to note the correlation of the authors' life data on polyethylene terephthalate (Mylar), to a 50-per-cent DS end point, with data obtained at the Westinghouse Research Laboratories using a quite different test. In this test the Mylar film was aged to an end point in a mechanical test. Failure was considered to occur at the time the film cracked when a strip of it was bent back on itself at 180 degrees. This is an even more arbitrary end point than the 50-per-cent DS criterion. I am not prepared to justify it.

A comparison of the results on the two tests is shown in Fig. 8.

With regard to the calculation scheme in the authors' paper where they derive a curve of dielectric strength as a function of time, it should be thoroughly understood that the basis for this curve lies in the assumption of a first-order reaction rate dependence of DS on time. The authors do not seem to have justified this. It should be established by experimental data at least at one temperature, i.e., by a plot of the logarithm of the DS versus time, which should prove to be linear.

As much faith as I have in this chemical rate analogy for insulation aging, I feel that we can assume more than is justified out of the mathematics in some cases. Insulation deterioration is frequently such a complicated process that the simpler rules will not apply. The validity of their application should be tried in each new situation.

L. C. Whitman and Paul Doigan: Dr. Dakin is correct in his contention that for strict functional evaluation a criterion of insulation aging other than dielectric strength may be proper. For example, Mr. Moses[1] has classified three distinct uses of insulation, namely a dielectric barrier, physical separation, and physical support. However, as a measure of rate of deterioration the DS criterion may still be valid in all three cases. These data then need only conversion to the proper scale of time and end point to be of value as a measure of the aging of significant properties in these other cases. The correlation of Dr. Dakin's mechanical aging data on Mylar with the dielectric strength criterion data strongly bears this out. The coincidence of end points is, of course, fortuitous but the rate of deterioration is evidently correctly measured by either method.

The use of the 50-per-cent dielectric strength criterion has been generally accepted for use on dielectric barrier material. For other situations some other per-cent criterion may be justifiable. One aim of this article was to permit the calculation of other percentage criteria from data experimentally obtained using a 50-per-cent criterion.

The assumption of a first-order reaction rate dependence of DS as a function of time is basic to the methods of calculations made in the article. This may not always be true and should be checked when this method is used.

For asbestos impregnated with a silicone some data have been presented by Dr. Dakin and others. These data approximate a straight line if plotted using logarithm of breakdown voltage as a function of time at a constant temperature.

The authors' data on Mylar at a constant temperature is not yet of sufficient time duration completely to justify a linear relationship. However, in this case the application of the theory developed gave a good check on test points made at a different per-cent voltage criterion as was shown in the article. This indicates that the straight-line relationship between logarithm of dielectric strength and time is also true for Mylar.

REFERENCE

1. THE RE-EXAMINATION OF TEMPERATURE STANDARDS FOR ELECTRICAL INSULATION, Graham Lee Moses. *AIEE Transactions*, vol. 71, pt. III, Aug. 1952, pp. 681–85.

Rectifier Arc-Back Study on the Analogue Computer

J. K. DILLARD
MEMBER AIEE

C. J. BALDWIN, JR.
ASSOCIATE MEMBER AIEE

ONE OF the most difficult problems encountered in the application of power rectifiers is the calculation of arc-back currents. Arc back is the term which refers to conduction of current through the rectifier in the reverse direction, i.e., from cathode to anode. It provides, simultaneously, a short circuit on the rectifier transformer low-voltage winding and paths for the flow of direct current from the d-c bus. Consequently, arc-back currents are nearly always higher than currents resulting from a 3-phase short circuit on the rectifier transformer. Knowledge of the magnitude of arc-back current is important because the resulting stresses may damage the rectifier or associated transformer and switchgear.[1–3] The magnitude of arc-back current can be obtained from the curves presented in this paper without restoring to the use of complicated mathematics. In addition, the curves provide a way to evaluate readily the effect of a change in circuit constants, a consideration which is important when reactors must be installed to limit stresses on the equipment.

Improvements in rectifier design, principally in design of tubes, have led to a continual decrease in arc-back rate of rectifiers over the years. The benefit of smaller numbers of arc backs has been offset, however, by higher arc-back currents which result from the tendency to operate more and more rectifiers in parallel at higher voltages. On modern large electrochemical potlines arc-back currents approach 100,000 amperes. These high currents emphasize the significance of arc-back calculations to designers of rectifier equipment.

Analytical determination of arc-back currents from the differential equations describing rectifier operation is a long and complicated procedure because several different modes of operation take place as other tubes on the faulted winding of the transformer start and stop conduction. For this reason, a great deal of attention has been given to the development of approximate methods for estimating the magnitudes of fault currents.[4–6] None of these approximate methods has proved wholly satisfactory, the results being much too conservative. There is a growing practice to temper these results by applying an arbitrary reduction factor based on experience. This is not desirable, either, because of the lack of sufficient field data. Since it is no simple matter to make arc-back tests in the field, these data are not likely to be available for some time.

This paper presents empirical curves showing the relationship between circuit constants and the first crest of arc-back current in the faulty anode and in the cathode circuit of a double-wye rectifier. Limiting action of the circuit breaker is neglected. The current at the first crest is of interest because it occurs about the same time that circuit-breaker action starts to limit the fault current. If the circuit breaker functions properly, the maximum value of fault current does not exceed the first crest value. The curves were plotted from the data obtained by setting the rectifier circuit up in miniature on the analogue computer. Miniature-system methods have been used for a long time to study power-system problems where analytical calculation is difficult or impractical. Apparently the convenience of miniature systems in rectifier studies has been overlooked until recently. However, these techniques do have interesting possibilities which have been demonstrated in two studies of rectifier voltage regulation as well as in this paper.[7,8]

Circuit Behavior

To provide a fundamentally sound basis for the general study of the arc-back problem, several items of circuit behavior under arc-back conditions must be considered. Included are current contributions from both a-c and d-c sources, drop in the d-c bus voltage, contribution from the unfaulted wye, and performance of the interphase transformer. The circuit diagram for the common double-wye connection is shown in Fig. 1. For large capacity installations the high-voltage winding of the rectifier transformer is fed from a low-impedance 3-phase source of power. When an arc back occurs on one of the ignitrons, fault current is fed from the 3-phase source through the low system impedance in series with the trans-

---

Paper 54-43, recommended by the AIEE Electronic Power Converters Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 16, 1953; made available for printing November 17, 1953.

J. K. DILLARD and C. J. BALDWIN, JR., are with Westinghouse Electric Corporation, East Pittsburgh, Pa.

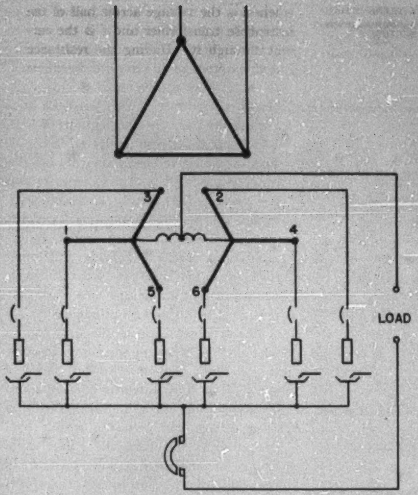


Fig. 1. Six-phase delta double-wye connection

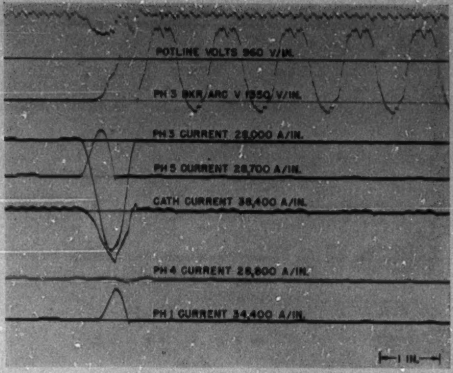


Fig. 2. Artificial arc-back test on one unit of a 10,000-kva quadruple-wye rectifier at a large electrochemical installation

former impedance to the fault. This is not, however, the only component of fault current flowing as a result of the arc back. If the load on the rectifier d-c bus is rotating machinery, rotor inertia will cause continued generation of a counterelectromotive force for some time after the arc back occurs. This generated direct voltage will cause a current to flow through the machine impedances, and bus voltage will be maintained at a value equal to the generated voltage less the impedance drop. Even in installations with no rotating machinery, current can be fed from the d-c bus. If there are parallel rectifiers on the bus, the faulty rectifier will appear as additional load, and parallel units will tend to keep bus voltage up at the terminals of the faulty unit.

The current contribution from the d-c bus depends upon the effective value of bus voltage during the arc back. This in turn is determined by either rotor inertia and internal drop of rotating machinery or regulation of parallel units on the bus. To predict the fault current accurately at first crest after occurrence of an arc back, the effective value of d-c bus voltage during this period must be known. Field data have shown that the bus voltage does not remain at 100 per cent of its normal value for as long as a half-cycle, but rather drops discernibly even with a number of parallel units. Fig. 2 shows oscillograms taken at a large aluminum reduction plant during artificial arc-back tests on one unit of a 10,000-kva quadruple-wye rectifier. For the test illustrated there were two quadruple-wye units on the bus, and the bus voltage at crest current was approximately 65-per-cent normal. Other tests with more units in parallel showed drops to about 75 or 80 per cent of normal voltage. These figures are believed to be slightly low because of the laminated bus construction on which the measurements were taken. Bus voltage at the terminals of the rectifier under test included the drop in the bus structure from parallel units, each of which is connected to a different bus lamination. Since each lamination carried a different current, potline voltages measured between different laminations of the bus would differ in magnitude. Based on these data a conservative assumption is that the d-c bus voltage drops at least 10 per cent even with a large number of

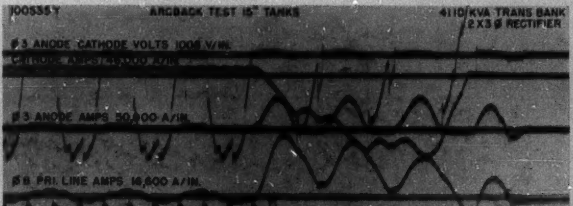


Fig. 3. Laboratory arc-back test on a 4,110-kva double-wye rectifier

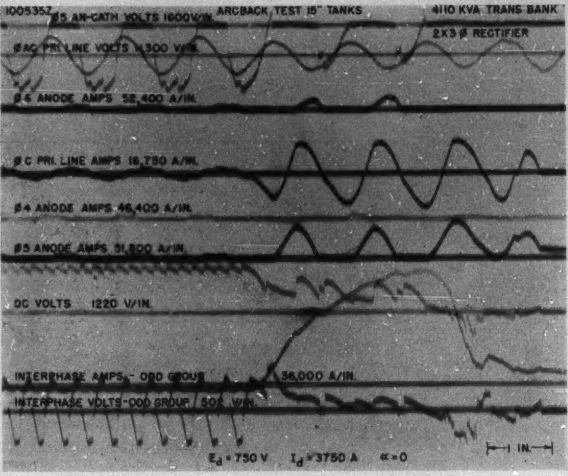


Fig. 4. Laboratory arc-back test on a 4,110-kva double-wye rectifier

parallel units on the bus; the same assumption is not unreasonable for rotating machinery since there is a voltage drop caused by fault current flowing through the internal machine impedance.

Another item of circuit behavior meriting consideration is the performance of the unfaulted wye group in the same double-wye with the faulty ignitron. Previously published oscillograms have shown very little current contribution from the unfaulted wye during the first half-cycle.[5] This is confirmed in Figs. 3 and 4, which are oscillograms taken in the laboratory on a 4,110-kva rectifier. Compared to the currents in anodes 3 and 5, those of anodes 2, 4, and 6 are insignificant. If the contribution from the unfaulted wye is very small, the wye group can be neglected in the mathematical analysis, and the problem thereby greatly simplified.

Interphase Transformer

While the unfaulted wye can be neglected, that half of the interphase transformer through which the fault current in the d-c path is flowing should not be neglected. There has been a great deal of speculation as to the performance of the interphase transformer during an arc back, especially concerning the impedance it presents to the fault current. Considering half of the interphase transformer as some unknown inductance L_t and resistance R_t in the d-c circuit, a differential equation can be written

$$e = iR_t + L_t\frac{di}{dt} \qquad (1)$$

where e is the voltage across half of the interphase transformer and i is the current through it. Having the resistance and the voltage and current as functions of time, the equivalent inductance L_t can be determined. This has been done graphically from the oscillograms of Figs. 3 and 4, and the results plotted in Fig. 5. In what is believed to be typical performance, the interphase transformer in this test saturated erratically in the first half-cycle, approaching saturation near the time of the first current crest. Calculated inductance of half of this transformer fully saturated is about 10 microhenrys, while apparent inductance at the first current crest was about 35 microhenrys.

The peculiar shape of the curve can be explained in terms of phenomena occurring in both the unfaulted and faulted wyes. The first discontinuity, at about 56 degrees, occurs when the ignitron on phase 2 stops conduction of normal load current. At this time there is a rapid change of current, i.e., a high $d\Phi/dt$, in the half of the interphase transform associated with the even group. Since the flux in the core is common to both

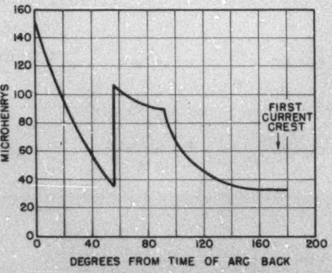


Fig. 5 (left). Apparent inductance of half of the interphase transformer during laboratory arc-back test

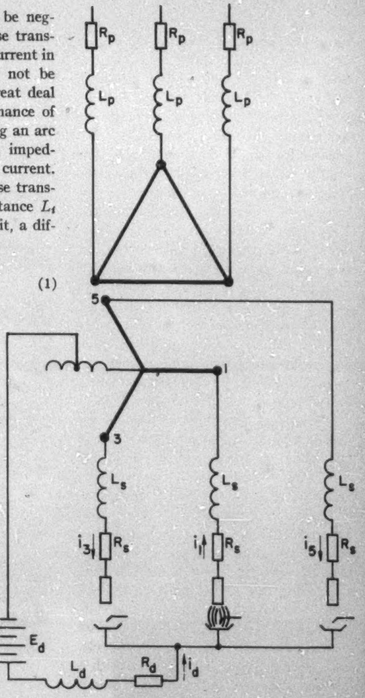


Fig. 6 (right). Equivalent circuit for mathematical analysis

Nomenclature

Current directions and impedance locations are shown in Fig. 6.

$\mathbf{a}$ = vector operator $1\underline{/120°}$
E_d = effective d-c bus voltage
E_s = rated line-to-neutral voltage of the transformer low-voltage winding
i_{ac} = current in tube 1 due to the alternating voltage only
i_{a1} = positive-sequence current
i_{a2} = negative-sequence current
i_{a0} = zero-sequence current
i_d = current in the d-c loop circuit
i_{dc} = current in the d-c loop circuit due to the direct voltage only
i_{d1} = current in the d-c loop circuit at the beginning of interval 1 = negative of normal load current
i_{d2} = current in the d-c loop circuit at the beginning of interval 2
i_{d3} = current in the d-c loop circuit at the beginning of interval 3
i_1 = current in tube 1. The second subscript indicates initial currents at the beginning of the respective intervals
i_3 = current in tube 3. The second subscript indicates initial currents at the beginning of the respective intervals
K = positive- and negative-sequence offset factor
K_0 = zero-sequence offset factor
L_d = inductance in the d-c circuit (the d-c loop and the effective inductance of half of the interphase transformer)
L_e = total per phase inductance presented to the fault (anode leads, transformer, high-voltage-side reactors, and a-c system)
L_p = inductance on the high-voltage side of the transformer (high-voltage-side reactors and a-c system but not the transformer)
L_s = inductance on the low-voltage side (anode leads and transformer)
L_t = line-to-neutral commutating inductance of the transformer only
$L_1=\frac{L_p}{6}+\frac{L_s}{2}+L_d$
$L_2=\frac{L_s}{3}+L_d$
$p=d/dt$
R_a = equivalent resistance representing arc drop
R_d = resistance in the d-c circuit (the d-c loop and half of the interphase transformer)
R_e = total per phase resistance presented to the fault (arc-drop resistance R_a and anode leads, transformer, reactors, and a-c system)
R_p = resistance on the high-voltage side of the transformer (high-voltage-side reactors and a-c system but not the transformer)
R_s = resistance on the low-voltage side (arc-drop resistance R_a, anode leads, and transformer)
R_t = line-to-neutral commutating resistance of the transformer only
$R_1=\frac{R_p}{6}+\frac{R_s}{2}+R_d$
$R_2=\frac{R_s}{3}+R_d$
t = time
t_1 = beginning of arc back as measured from time reference
t_2 = beginning of interval 2
t_3 = beginning of interval 3
$X=2\pi fL$
$Z_d=\sqrt{R_d^2+X_d^2}$
$Z_e=\sqrt{R_e^2+X_e^2}$
$Z_s=\sqrt{R_s^2+X_s^2}$
$Z_1=\sqrt{R_1^2+X_1^2}$
$\theta=\tan^{-1}\frac{X_e}{R_e}$
$\theta_{d1}=\tan^{-1}\frac{X_1}{R_1}$
$\omega=2\pi f$ = angular frequency

---

halves, there results a high induced voltage in the side under investigation, and the terminal voltage and apparent inductance jump to higher values. Likewise, the second distinct break in the curve, at 92 degrees, occurs when the ignitron on phase 5 begins conduction of fault current. The polarity is such as to cause increased saturation of the interphase transformer.

From the plot of Fig. 5, there is apparently no single value of inductance which can be used to calculate the arc-back current accurately. A conservative assumption must be made to permit mathematical treatment of the problem.

Mathematical Development

Based on the assumptions already discussed, the double-wye circuit of Fig. 1 can be reduced to the circuit of Fig. 6 for mathematical analysis. The solutions to the differential equations for the case of all secondary impedance have been presented in an earlier article[5]; however, with the increasing use of current-limiting reactors in the high-voltage lines, the influence of primary reactance on arc-back currents should be considered. The equations of the Herskind-Schmidt-Rettig article[5] have been extended to include primary quantities by redefining the constants L_1, L_2, R_1, and R_2 to include primary impedances. In addition, algebraic signs appearing in front of the second terms of equations 2 and 7 have been corrected. To describe circuit action to the first crest of fault current, only the equations of the first three intervals are needed. During interval 1, tubes 1 and 3 are conducting. When the anode-to-cathode voltage on tube 5 becomes zero tube 5 picks up, and interval 2 begins, lasting until tube 3 stops conduction. Then, during interval 3, only tubes 1 and 5 are conducting. Time reference is the voltage zero on phase 3.

Interval 1 – t_1 to t_2

$$i_d=\frac{E_d}{R_1}\left[1-\epsilon^{-\frac{R_1}{L_1}(t-t_1)}\right]-\frac{\sqrt{2}E_s}{2Z_1}\left[\sin\left(\omega t+\frac{\pi}{3}-\theta_{d1}\right)-\epsilon^{-\frac{R_1}{L_1}(t-t_1)}\sin\left(\omega t_1+\frac{\pi}{3}-\theta_{d1}\right)\right]+i_{d1}\epsilon^{-\frac{R_1}{L_1}(t-t_1)} \tag{2}$$

$$i_1=\frac{\sqrt{6}E_s}{2Z_e}\left[\sin\left(\omega t-\frac{\pi}{6}-\theta\right)-\epsilon^{-\frac{R_e}{L_e}(t-t_1)}\sin\left(\omega t_1-\frac{\pi}{6}-\theta\right)\right]+\frac{1}{2}i_d+\left(i_{11}-\frac{1}{2}i_{d1}\right)\epsilon^{-\frac{R_e}{L_e}(t-t_1)} \tag{3}$$

Interval 2 – t_2 to t_3

$$i_d=\frac{E_d}{R_2}\left[1-\epsilon^{-\frac{R_2}{L_2}(t-t_2)}\right]+i_{d2}\epsilon^{-\frac{R_2}{L_2}(t-t_2)} \tag{4}$$

$$i_1=\frac{\sqrt{2}E_s}{Z_e}\left[\sin\left(\omega t-\frac{\pi}{3}-\theta\right)-\epsilon^{-\frac{R_e}{L_e}(t-t_2)}\times\sin\left(\omega t_2-\frac{\pi}{3}-\theta\right)\right]+\frac{1}{3}i_d+\left(i_{12}-\frac{1}{3}i_{d2}\right)\epsilon^{-\frac{R_e}{L_e}(t-t_2)} \tag{5}$$

$$i_3=\frac{\sqrt{2}E_s}{Z_e}\left[\sin(\omega t-\theta)-\epsilon^{-\frac{R_e}{L_e}(t-t_2)}\times\sin(\omega t_2-\theta)\right]-\frac{1}{3}i_d+\left(i_{32}+\frac{1}{3}i_{d2}\right)\epsilon^{-\frac{R_e}{L_e}(t-t_2)} \tag{6}$$

Interval 3 – t_3 to t_4

$$i_d=\frac{E_d}{R_1}\left[1-\epsilon^{-\frac{R_1}{L_1}(t-t_3)}\right]+\frac{\sqrt{2}E_s}{2Z_1}\times\left[\sin(\omega t-\theta_{d1})-\epsilon^{-\frac{R_1}{L_1}(t-t_3)}\times\sin(\omega t_3-\theta_{d1})\right]+i_{d3}\epsilon^{-\frac{R_1}{L_1}(t-t_3)} \tag{7}$$

$$i_1=-\frac{\sqrt{6}E_s}{2Z_e}\left[\cos(\omega t-\theta)-\epsilon^{-\frac{R_e}{L_e}(t-t_3)}\times\cos(\omega t_3-\theta)\right]+\frac{1}{2}i_d+\left(i_{13}-\frac{1}{2}i_{d3}\right)\epsilon^{-\frac{R_e}{L_e}(t-t_3)} \tag{8}$$

The condition for the end of interval 1 is zero anode-to-cathode voltage on tube 5, assuming free conduction. In a step-by-step calculation this condition for the end of interval 1 must be satisfied, as expressed by equation 9

$$i_d\left(\frac{R_p}{3}-R_d\right)+pi_d\left(\frac{L_p}{3}-L_d\right)+i_1R_s+pi_1L_s=\sqrt{2}E_s\sin\left(\omega t-\frac{2\pi}{3}\right) \tag{9}$$

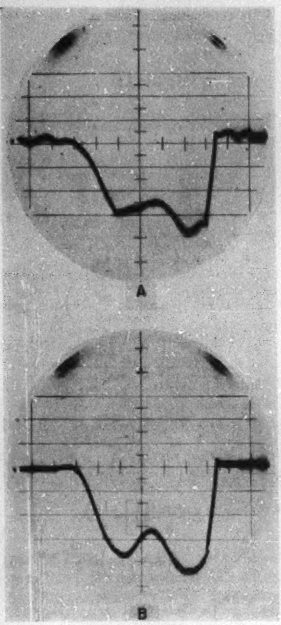


Fig. 7. Computer oscillograms for the germanium analogue

A—D-c loop current
B—Anode current in the faulty tube

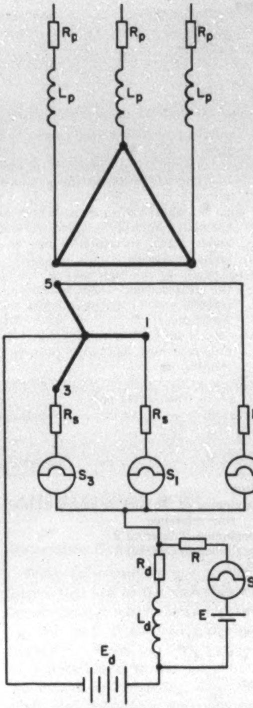


Fig. 8. Computer circuit for switch analogue

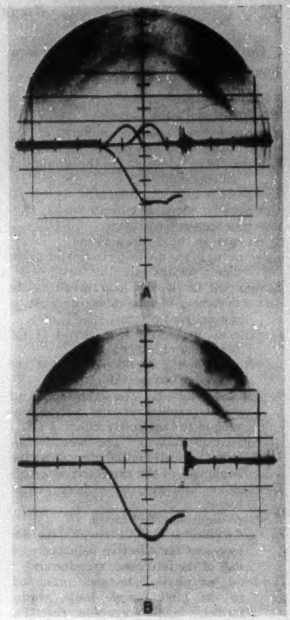


Fig. 9. Computer oscillograms for the switch analogue

A—D-c loop current and anode currents in the unfaulted tubes
B—Anode current in the faulty tube

The quantities pi_d and pi_1 can be obtained by differentiating equations 2 and 3. The condition for the end of interval 2 is $i_3=0$ in equation 6.

As in previous work, arc drop of the ignitron is represented by an equivalent resistance. The value of the resistance is calculated from tube characteristics, and, if desirable, the circuit-breaker arc drop can be included in a similar manner.

Equations 2 through 8 can be derived using either the differential equations with the Laplace transformation or the method of symmetrical components. Analysis by symmetrical components provides a simple description of circuit performance in terms familiar to power-system engineers. The arc back can be considered as a series of secondary faults through an impedance, and the contribution from the d-c bus voltage obtained by using the principle of superposition. With the direct voltage short-circuited, the arc back appears as a double-line-to-ground fault through the impedance Z_d during the first interval. When tube 5 begins conduction, current resulting from the alternating voltage is given by equations for a 3-phase fault. Again in the third interval the arc back appears as a double-line-to-ground fault through an impedance. The transient parts of the solution can be added to the symmetrical-solution by applying the proper decrement to the initial values. With the alternating voltage short-circuited, the contribution from the direct voltage is easily obtained as a d-c transient.

The viewpoint afforded by symmetrical components is especially valuable in comparing the effects of primary and secondary impedances. For a fault on the low-voltage side of a delta-wye step-down transformer, the zero-sequence impedance includes only the transformer impedance and those impedances on the wye side of the transformer. Since the arc back actually involves ground faults in certain intervals, adding primary reactance does not have the same effect as secondary reactance in limiting fault current. Current-limiting reactors placed in the high-voltage lines do not add to the zero-sequence impedance.

Symmetrical components offer a simple check on the transform solutions in determining the impedance presented to the fault current in the different intervals. An example is presented in the Appendix in which the impedance presented to the ground current, i.e., current in the d-c circuit, is derived for the intervals during which two tubes conduct.

Analogue Computer Study

Transient phenomena in networks where analytical methods are either hopelessly complicated or extremely time-consuming are best handled on the analogue computer.[9] The network is constructed in miniature with precision circuit elements and supplied by a low-voltage source. Synchronous switches are used to apply the transient repetitively

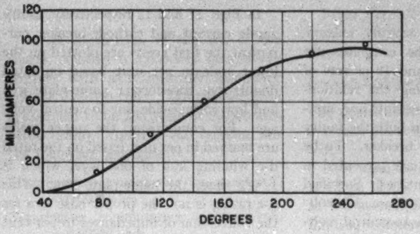


Fig. 10. Comparison of computer results and calculated curve of anode current in the faulty tube

— Calculated curve
0 Computer points

so that the solution appears as a steady picture on the oscilloscope. In effect, oscilloscope measurements are then exact solutions to the differential equations describing the transient. Since circuit elements can be varied over a wide range, this treatment lends itself to the rapid study of many variables.

In a study concerned with the first-crest arc-back current in a double-wye rectifier, the circuit of Fig. 6 is adequate to describe circuit behavior. Several possible ways of representing this circuit on the computer were investigated. The use of thyratrons for an ignitron analogue was dismissed because of their high tube drop compared to safe computer supply voltages. Selected selenium cells with very low forward drop, which have been used successfully in the past,[7,8] gave transient difficulties because of their capacitance. In an attempt to obtain both low forward drop and negligible capacitive effects, germanium diodes were used for the ignitron representation. While the forward germanium drop is somewhat higher than that of selenium, it still permitted the use of a 22-volt power supply. With germanium cells representing ignitrons, the arc back could be applied across the rectifier of one phase with: 1. a short-circuiting synchronous switch; 2. a short-circuiting switch in series with a resistor representing arc drop; or 3. a short-circuiting switch in series with reversed diodes. In the preliminary study method 2 was used and oscillograms were obtained from the computer as shown in Fig. 7. Fig. 7(A) shows the current in the d-c loop, i.e., the cathode-breaker current; Fig. 7(B) shows the arc-back current in the faulty anode. The synchronous switch removed the transient a short time after the second current crest. The shapes of these current waves should be compared with those of Figs. 2, 3, and 4, which were taken on actual double-wye or quadruple-wye rectifiers.

The germanium analogue is adequate for the study of a particular system in which the ignitron arc-drop characteristic can be matched by a judicious combination of the germanium characteristic and a series anode resistor. However, it is not adaptable to the gathering of general data which can be presented in per-unit quantities and which is applicable to any arc-drop characteristic. To obtain such data, it was necessary to resort to a strict representation of the circuit from which equations 2 through 9 were derived, i.e., a circuit with arc drop represented by an equivalent resistance only.

The analogue used in the final study is illustrated in Fig. 8. Synchronous switches with series resistors were used to represent the ignitrons and were adjusted to operate in proper sequence. Since the circuit was not operating as a rectifier before the transient was initiated, an additional switch and battery were provided to give the initial load current in the d-c inductance. Oscilloscope observations showed the effect from initial currents to be negligible, and they could have been omitted with little consequence. The d-c bus voltage E_d, represented by a battery, was held to approximately 90 per cent of normal bus voltage. During each repetitive sequence, switch S_d closed first on the de-energized circuit to provide the initial load current in inductance L_d. At the time giving a maximum fault current, switches S_1 and S_3 closed to initiate the arc back on phase 1. Simultaneously switch S_d opened. When the voltage across switch S_5, i.e., the anode-to-cathode voltage, went through zero, switch S_5 closed to begin the second interval. The opening of switch S_3 marked the end of the second interval and was adjusted to take place when the current through it reached zero. After the crest arc-back current occurred, switch S_5 was opened at its current zero. Switch S_1 was then opened to de-energize the circuit. With each change of circuit constants the switches were adjusted to operate at the proper time, and the crest arc-back currents were read from the oscilloscope.

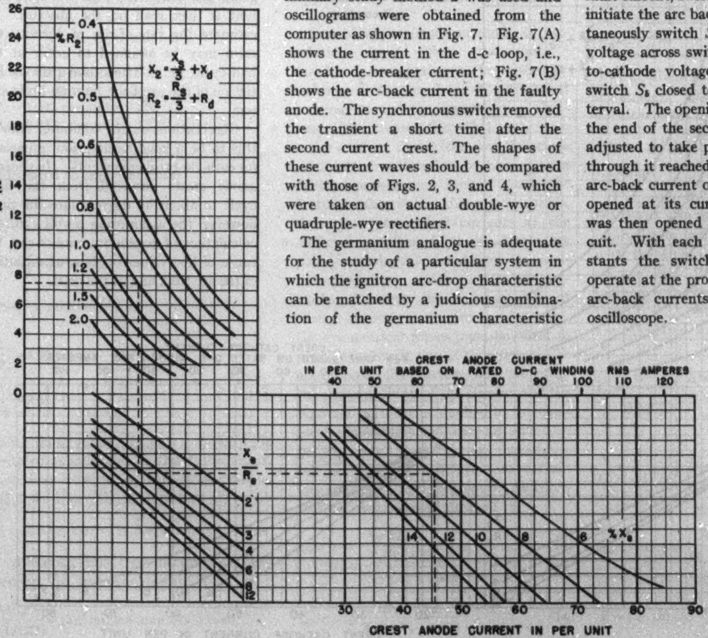


Fig. 11. First-crest anode current in the faulty tube as a function of circuit constants

Typical oscillograms obtained from the switch analogue are shown in Fig. 9. Fig. 9(A) shows the currents flowing in the normal anodes above the abscissa and the d-c loop current below the abscissa. Current in phase 3 is, of course, on the left since that phase is just beginning conduction when the arc back occurs. Fig. 9(B) shows the arc-back current in the faulty anode, the sum of the currents shown in Fig. 9(A). The oscillograms of Fig. 9 were not made with the same combination of circuit constants used for Fig. 7.

To verify the accuracy of the analogue representation, a typical case was set up on the computer, and the currents were calculated using the exact equations 2 through 9. The verification is shown in Fig. 10, in which the solid curve is the calculated anode current; test points from the oscilloscope are shown in small circles. The ordinate of the curve is computer milliamperes, while the abscissa is degrees of time measured from the time reference of zero transformer voltage on phase 3. Magnitude and wave shape of the d-c loop current were also verified.

Use of the Charts

Data taken from the computer are presented in Figs. 11 and 12 as sets of empirical curves showing the relationship between circuit constants and first-crest arc-back current in both the faulty anode and the cathode breaker. To be quite general, the data are presented in per-unit or per-cent quantities. Specified base quantities are 3-phase kilovolt-amperes (kva) and line-to-neutral voltage in kilovolts. In terms of these base quantities, line-to-neutral per-unit impedance can be calculated

$$Z \text{ in per unit} = \frac{(Z \text{ in ohms})(\text{3-phase base kva})}{3{,}000\ (\text{line-to-neutral base kv})^2} \tag{10}$$

Base phase current and base line-to-neutral ohms are calculated from equations 11 and 12

$$\text{Base current} = \frac{\text{3-phase base kva}}{3\ (\text{line-to-neutral base kv})} \tag{11}$$

$$\text{Base ohms} = \frac{\text{line-to-neutral base volts}}{\text{base current}} \tag{12}$$

In Figs. 11 and 12 respectively, faulty anode current and cathode-breaker current at the first crests are plotted on the lower scale in per unit based on rated double-wye transformer name-plate kva and low-voltage-side line-to-neutral voltage. For comparison, the upper scales are marked in per unit based on the rated d-c winding kva of one wye, which is $1/\sqrt{2}$ times the name-plate kva. This d-c rating is not the proper base kva for the conversion of impedances to per-cent quantities. To apply the charts to a quadruple-wye connection, the circuit can be considered as two double-wye transformers in parallel. The name-plate commutating ohms therefore must be doubled before being put in per cent on the double-wye kva base, i.e., a base numerically equal to half the quadruple-wye name-plate kva. To use the charts for either circuit, all circuit constants must be in per cent on the same base quantities.

The charts are entered from the left margin at the given X_2/R_2 ratio and this point projected horizontally to the given per cent R_2. The intersection is projected vertically downward to the given X_e/R_e ratio and the new intersection projected horizontally to the right to the given per cent X_e. The first-crest current is then read on the abscissa. To estimate the rate-of-rise of current, the crest may be assumed to occur at the half-cycle, i.e., 1/120 second after the initiation of the arc back on a 60-cycle system.

Observations on the computer showed this to be a good assumption, especially for cathode-breaker currents. In general, anode crests occurred from the half-cycle to about 40 degrees later in time; however, the assumption is still valid for calculation of maximum rate-of-rise because of low values of slope at initiation of the arc back and at the crest.

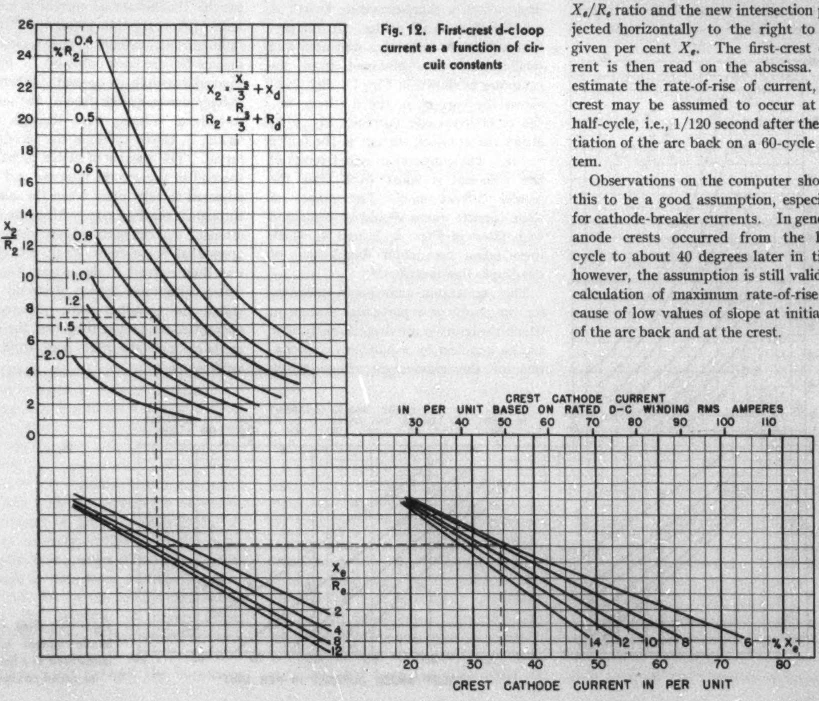


Fig. 12. First-crest d-c loop current as a function of circuit constants

Example

As an illustration of the use of the charts, consider a 5,000-kva double-wye rectifier with a transformer secondary voltage $E_s = 650$ volts. Ignitron arc drop is assumed to be 203 volts at 80,000 amperes (amp). Assume the rectifier has the following constants:

Transformer commutating reactance $= X_t = 0.0203$ ohm
Transformer commutating resistance $= R_t = 0.00203$ ohm
Equivalent arc-drop resistance of the ignitron $= R_a = 203/80{,}000 = 0.00254$ ohm
Interphase and d-c loop resistance $= R_d = 0.000507$ ohm
Interphase and d-c loop reactance $= X_d = 22.4$ microhenrys $= 0.00845$ ohm

For this example the rectifier has no current-limiting reactors; reactances of the a-c system and the anode leads are negligible. First, the rated base current and base ohms must be calculated from equations 11 and 12

$$\text{Base current} = \frac{5{,}000}{3(0.650)} = 2{,}560 \text{ amp}$$

$$\text{Base ohms} = \frac{650}{2{,}560} = 0.254 \text{ ohm}$$

The impedances then must be put in per cent (%) on these base quantities

$$X_t = \frac{0.0203}{0.254} = 0.08 \text{ per unit} = 8\%$$

$R_t = 0.008$ per unit $= 0.8\%$
$R_a = 0.01$ per unit $= 1\%$
$R_d = 0.002$ per unit $= 0.2\%$
$X_d = 0.0333$ per unit $= 3.33\%$

X_e, R_e, X_2, R_2, and the ratios must be calculated to enter the chart

$$X_e = 8\%$$

$$R_e = 0.8 + 1.0 = 1.8\%$$

$$X_2 = \frac{8}{3} + 3.33 = 6\%$$

$$R_2 = \frac{1.8}{3} + 0.2 = 0.8\%$$

$$\frac{X_e}{R_e} = 4.45$$

$$\frac{X_2}{R_2} = 7.5$$

With these values, the chart of Fig. 11 is entered as indicated by the dashed lines, and first-crest anode current is read as 45.5 per unit. Or, in amperes

$$i_1 = 45.5\,(2{,}560) = 116{,}500 \text{ amp}$$

Assuming this crest occurs at 1/120 second, the approximate rate-of-rise is

$$\frac{116{,}500}{(1/120)} = 13.9 \text{ million amp per second}$$

From Fig. 12 the first-crest cathode-breaker current is 34.5 per unit. In amperes

$$i_d = 34.5(2{,}560) = 88{,}400 \text{ amp}$$

The approximate rate-of-rise is 10.6 million amp per second for the cathode current.

Now assume that 4-per-cent reactance is added in the high-voltage lines either as a-c system reactance or as current-limiting reactors. Since this does not change X_s or R_s, X_2 and R_2 are unchanged. Assume the reactor has 0.4-per-cent resistance

$$R_e = 1.0 + 0.8 + 0.4 = 2.2\%$$

$$X_e = 8 + 4 = 12\%$$

$$\frac{X_e}{R_e} = 5.45$$

Entering the chart of Fig. 11 again with the same X_2 and R_2, the crest anode current is read as 36.5 per unit, or 93,500 amp. The approximate rate-of-rise is 11.2 million amp per second. In a similar manner the cathode current is 30.5 per unit or 78,000 amp. The approximate rate-of-rise of cathode current is 9.36 million amp per second.

Approximate Method of Calculation

Several approximate methods of calculation have been presented.[4-6] Another approximation can be applied by making use of the methods of symmetrical components and the superposition theorem. The assumption that only two tubes conduct during the first half-cycle permits the problem to be treated as a double-line-to-ground fault through the fault impedance Z_d. Using components, the crest symmetrical currents at the half-cycle resulting from the alternating voltage can be calculated from the positive-, negative-, and zero-sequence networks. Because the currents cannot change instantly in the circuit inductance, the symmetrical waves must be offset to provide a current zero at time zero. The appropriate offset factors, determined by the phase position of the current and the X/R ratios, must be applied to the component currents before they are added together. For simplification, resistance can be neglected in determining current magnitudes. This procedure leads directly to equations 13 through 18

$$i_{a1} = \frac{\sqrt{2}E_s}{X_e + \dfrac{X_e(X_s + 3X_d)}{X_e + (X_s + 3X_d)}} \tag{13}$$

$$i_{a2} = -i_{a1}\frac{X_s + 3X_d}{X_e + (X_s + 3X_d)} \tag{14}$$

$$i_{a0} = -i_{a1}\frac{X_e}{X_e + (X_s + 3X_d)} \tag{15}$$

The offset factor to be applied to i_{a1} and i_{a2} is very nearly 2.0 for most practical cases. It can be approximated by equation 16

$$K = 1 + \sin\tan^{-1}\frac{X_e}{R_e} \tag{16}$$

Because of the phase position of the zero-sequence current, its offset factor is very nearly zero, as given by equation 17

$$K_0 = \cos\tan^{-1}\frac{X_1}{R_1} \tag{17}$$

The current in the faulty anode at the half-cycle resulting from the alternating voltage is obtained by the usual addition of component currents with the vector operator **a**

$$i_{ac} = \left|\mathbf{a}^2 K i_{a1} + \mathbf{a} K i_{a2} + K_0 i_{a0}\right| \tag{18}$$

With the alternating voltage short-circuited, the current flowing in the d-c loop as a result of the bus voltage is a simple d-c transient

$$i_{dc} = \frac{E_d}{R_1}\left[1 - \epsilon^{-\frac{R_1}{L_1}t}\right] \tag{19}$$

If i_{dc} is evaluated at the half-cycle, the approximate crest current in the faulty anode is, by superposition

$$i_1 = i_{ac} + \frac{1}{2}i_{dc} \tag{20}$$

since i_{dc} divides between the two conducting anodes. The current in the d-c loop circuit is given by equation 21

$$i_d = i_{dc} - 3K_0 i_{a0} \tag{21}$$

Example

The numerical example making use of the charts also can be calculated with the approximate method. Consider the case with the 4-per-cent reactor in the high-voltage lines. Using the approximate method with ohmic values, and the relation $X_s + 3X_d = 3X_2$, first-crest currents are calculated with equations 13 through 21

$$i_{a1} = \frac{\sqrt{2}(650)}{0.0305 + \dfrac{0.0305(0.0457)}{0.0762}} = 18{,}850 \text{ symmetrical amp}$$

$$i_{a2} = -18{,}850\left(\frac{0.0457}{0.0762}\right) = -11{,}300 \text{ symmetrical amp}$$

$$i_{a0} = -18{,}850\left(\frac{0.0305}{0.0762}\right) = -7{,}550 \text{ symmetrical amp}$$

$$K = 1 + \sin\tan^{-1}\frac{12\%}{2.2\%} = 1.982$$

Table I. Typical Constants for Large Electrochemical Installations

Constant	Minimum	Maximum
Rectifier transformer reactance, per cent	6	9
Rectifier transformer X/R ratio	8	11
Anode lead inductance, microhenrys	1	10
Anode lead resistance, ohms	0.0001	0.0010
Equivalent tube drop resistance, ohms	0.0018	0.0030
D-c loop inductance, microhenrys	15	40
D-c loop resistance, ohms	0.0003	0.0010
Interphase transformer, half-winding saturated inductance, microhenrys	2	20
Interphase transformer, half-winding resistance, ohms	0.0002	0.0004

$$X_1=\frac{4\%}{6}+\frac{8\%}{2}+3.33\%=8\%$$

$$R_1=\frac{0.4\%}{6}+\frac{1.8\%}{2}+0.2\%=1.16\%$$

$$K_0=\cos\tan^{-1}\frac{8\%}{1.167\%}=0.144$$

$$\begin{aligned}i_{ac}&=\left|1.982(18{,}850\underline{/240^\circ})+1.982\times\right.\\&\qquad\left.(-11{,}300\underline{/120^\circ})+0.144(-7550)\right|\\&=\left|-8590-j51{,}800\right|=52{,}500\text{ amp}\end{aligned}$$

If bus voltage is held to about 90 per cent of its normal value

$$i_{dc}=\frac{630}{0.00296}\left[1-\epsilon^{-\frac{0.00296\,(377)}{0.0203\,(120)}}\right]=78{,}500\text{ amp}$$

$$i_1=52{,}500+\frac{78{,}500}{2}=91{,}750\text{ amp}$$

$$i_d=78{,}500-3(0.144)(-7550)=81{,}760\text{ amp}$$

This is 98.2 per cent of the anode current obtained from the exact equations or Fig. 11, and the cathode current is 105 per cent of that obtained from Fig. 12. For reasonable X/R ratios, very little error would be introduced by assuming $K=2.0$ and $K_0=0$.

Both this method and those presented previously are approximations since they assume only two tubes conducting during the first half-cycle. The method of components does have the advantage for field calculations in that no special equations are needed. However, for the most accurate results the charts of Figs. 11 and 12 should be used.

Circuit Constants

Questions about estimated magnitudes of arc-back current often come up at rectifier stations for which exact values of all the circuit constants are unknown and cannot be determined readily. In these situations it is desirable to have at hand typical values of circuit resistance and inductance in order to obtain realistic estimates from the curves given in this paper. Most of the large rectifier installations in this country are employed in the electrochemical industry. Some idea of the range of constants from circuits used in this industry can be obtained from the values tabulated in Table I.

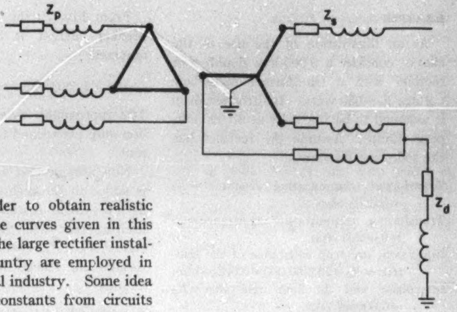


Fig. 13. Circuit diagram for double-line-to-ground fault through an impedance

Exact values of circuit constants on the d-c side of most large rectifier stations are difficult to obtain by calculation because of the complexity of the d-c bus structure. This is not true of the a-c side where impedances of the a-c supply and transformer can be determined quite accurately in most cases. But the resistances and inductances of the d-c circuit and interphase transformer and the arc drop across the rectifier at fault currents may have to be determined by field measurement. Much could be learned from a limited number of field measurements in the electrochemical industry since there is fair similarity between rectifier circuits in these large stations.

One other factor which enters the picture is the limiting action of the circuit breaker used to clear the arc back. A voltage opposing the build-up of fault current is inserted in the circuit as soon as the circuit breaker starts to open. This limiting action was not included in this study because the present practice is to apply switchgear on the basis of magnitudes and rates-of-rise of current which would exist if the circuit breaker were not in the circuit. Investigation of the effect of circuit-breaker arc voltage and determination of accurate values of circuit constants are, perhaps, the most fertile fields for further contribution to a better understanding of the arc-back problem.

Appendix. Effect of Primary and Secondary Reactance by the Method of Symmetrical Components

To derive the impedance presented to the current flowing in the d-c circuit during intervals 1 and 3, the arc back can be considered a double-line-to-ground fault through an impedance, as illustrated in Fig. 13. For this fault, the negative- and zero-sequence networks are paralleled and connected in series with the positive-sequence network. The rms current equations are

$$i_{a1}=\frac{E_s}{Z_{11}+\dfrac{Z_{22}(Z_{00}+3Z_f)}{Z_{22}+Z_{00}+3Z_f}}\tag{22}$$

$$i_{a0}=-i_{a1}\frac{Z_{22}}{Z_{22}+Z_{00}+3Z_f}\tag{23}$$

where Z_{11}, Z_{22}, and Z_{00} are positive-, negative-, and zero-sequence impedances respectively. Z_f is the fault impedance

For the delta-wye transformer connection, in terms of primary- and secondary-side impedances previously defined

$$\begin{aligned}&Z_{11}=Z_{22}=Z_p+Z_s\\&Z_{00}=Z_s\\&3Z_f=3Z_d\end{aligned}\tag{24}$$

Making these substitutions in equations 22 and 23

$$\begin{aligned}3i_{a0}&=-\frac{3E_s}{Z_{11}+2(Z_{00}+3Z_f)}\\&=-\frac{E_s}{2}\left[\frac{1}{\dfrac{Z_p}{6}+\dfrac{Z_s}{2}+Z_d}\right]\end{aligned}\tag{25}$$

Therefore, Z_1 of equations 2 and 7 is given by equation 26

$$Z_1=\frac{Z_p}{6}+\frac{Z_s}{2}+Z_d\tag{26}$$

References

1. Mercury Arc Power Rectifiers (book), O. K. Marti, H. Winograd. McGraw-Hill Book Company, Inc., New York, N. Y., 1930, pp. 35–43, 80–82, 188–91.

2. A New Multipole High-Speed Air Circuit Breaker for Mercury-Arc-Rectifier Anode Circuits and Its Relation to the Arc-Back Problem, J. W. Seaman, L. W. Morton. *AIEE Transactions (Electrical Engineering)*, vol. 61, Nov. 1942, pp. 788–96, 1057–61.

3. Arc-Backs in Rectifier Circuits—Artificial Arc-Back Tests, R. D. Evans, A. J. Maslin. *AIEE Transactions (Electrical Engineering)*, vol. 64, June 1945, pp. 303–11, discussion 439–41.

4. Calculation of the Backfiring Phenomenon in Rectifiers, H. Forssell. *Technical Achievements of ASEA Research*, ASEA, Vasteras, Sweden, 1946, pp. 58–64.

5. Rectifier Fault Currents—II, C. C. Herskind, A. Schmidt, Jr., C. E. Rettig. *AIEE Transactions*, vol. 68, pt. I, 1949, pp. 243–52.

6. PROTECTION OF ELECTRONIC POWER CONVERTERS, AIEE Committee Report. *AIEE Transactions*, vol. 69, pt. 1, 1950, pp. 813–29.

7. INFLUENCE OF A-C REACTANCE ON VOLTAGE REGULATION OF 6-PHASE RECTIFIERS, R. L. Witzke, J. V. Kresser, J. K. Dillard. *AIEE Transactions*, vol. 72, pt. I, July 1953, pp. 244–53.

8. VOLTAGE REGULATION OF 12-PHASE DOUBLE-WYE RECTIFIERS, R. L. Witzke, J. V. Kresser, J. K. Dillard. *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 689–97.

9. A LARGE-SCALE GENERAL-PURPOSE ELECTRIC ANALOGUE COMPUTER, E. L. Harder, G. D. McCann. *AIEE Transactions*, vol. 67, pt. I, 1948, pp. 664–73.

◆

Discussion

W. Fraser (Aluminum Laboratories Ltd., Arvida, Que., Canada): The analogue computer study described provides an accurate and rapid means of estimating rectifier fault currents. This method should provide valuable information to those concerned with the use and application of rectifier equipment.

The omission of the limiting action of the circuit breaker is fully justified. The rate of rise of current which would exist without this limiting action provides the switchgear designer with an indication of the duty requirements of d-c protective equipment

The statement made by the authors that "the tendency to operate more and more rectifiers in parallel . . . results in higher arc-back currents" is probably true in some cases; actually a greater contribution to higher arc-back currents is the use of individual rectifier units of higher current capacity.

H. Winograd (Allis-Chalmers Manufacturing Company, Milwaukee, Wis.): The authors are to be commended for an excellent job in applying the analogue computer to the study of rectifier circuit problems. They have used ingenuity in adapting the computer for this work and in the presentation of the results. Those of us who occasionally have to calculate the fault currents in rectifier circuits can best appreciate the value of the graphs in Figs. 11 and 12. I hope the authors will continue the good work.

The statement in the fourth paragraph of the paper that "Apparently the convenience of miniature systems in rectifier studies has been overlooked until recently" is not quite correct. We at Allis-Chalmers have utilized miniature rectifier circuits for this purpose on various occasions, employing thyratron tubes with small rectifier transformers to represent rectifier units. For example, the application of phase shifters for phase multiplication in the first large rectifier installation for an aluminum reduction plant, in 1938, was studied on a miniature setup before the phase shifters were installed in the aluminum plant. Five miniature 6-phase rectifier units with thyratron tubes were used to represent five 5,500-kw units supplying power to each of two potlines in the Alcoa, Tenn., plant of the Aluminum Company of America. The 30-phase system tried in miniature was applied to the large installation; it solved a serious telephone interference problem and set the pattern for phase multiplication in all subsequent large rectifier installations for electrochemical service. Some of the results obtained are given in reference 1 and a miniature test set is shown there in Fig. 4.

We have also used miniature rectifier systems with thyratron tubes for the study of electronic motor circuits, grid-control circuits, and other rectifier circuit problems. Miniature rectifier circuits with thyratron tubes require the use of higher voltages and power than the analogue computer and may not be as flexible in varying the circuit constants. However, they have advantages for studies involving phase control, and it is possible to obtain simultaneous oscillographic records in various parts of the rectifier circuit.

REFERENCE

1. WAVE SHAPE OF 30- AND 60-PHASE RECTIFIER GROUPS, O. K. Marti, T. A. Taylor. *AIEE Transactions (Electrical Engineering)*, vol. 59, April 1940, pp. 218–26.

A. Schmidt, Jr. (General Electric Company, Schenectady, N. Y.): The authors are to be congratulated for having tackled a difficult problem and arrived at a useful solution. A simplified method of calculation for general use has been needed for some time.

The authors have assumed the source of d-c feed as an infinite bus having a voltage equal to 90 per cent of the normal d-c bus voltage. It seems that a more realistic procedure would consist of adding the internal impedance of the source of d-c feed to the impedance of the d-c loop and using a higher value of voltage for the d-c feed. This value would be E_{d0} in case rectifiers are the source of d-c feed. If the rectifier load consists of d-c machines, their internal voltage would constitute the d-c feed voltage.

It is unfortunate that the authors did not present information on currents in the normal anodes. It is well known that a rectifier anode has a fairly definite fault current capacity[1] and that when this capacity is exceeded, as it may be when feeding into an arc back, consequential arc backs can be expected. If the currents in the normal anodes are known, the probability of consequential arc backs can be evaluated.

The figures given by the authors are somewhat pessimistic in two respects. In many cases two or more rectifiers are connected to a system. When one of these rectifiers arcs back, the feed of the parallel rectifier or rectifiers into the arc back reduces the transformer voltage of the faulty rectifier due to system impedance. This will reduce the contribution by the normal anodes in the wye which includes the faulty anode. Another pessimistic assumption, and one which has been made in previous studies, is the assumption that the second normal anode enters conduction at the time when its potential becomes positive. Most excitation systems are arranged so that the anode is not permitted to fire until a later time. If this later starting time were chosen, the magnitude of the fault current would be somewhat lower.

The authors have carried out the solution of equations given in a previous paper[2] to determine fault currents. This earlier paper also describes an approximate method of solution which has since been modified and used with results having a fair degree of accuracy. In applying this approximate method, the faulty rectifier is assumed to present to the source of d-c feed a counterelectromotive force having a-c and d-c components. These components are determined by inspection (see Fig. 7 of earlier paper[2]) to be approximately $0.25E_s$ for both the crest value of the a-c component and for the d-c component.

If this approximate method is applied to the authors' example, currents of 120,200 and 88,700 amperes are found for the faulty anode and the cathode respectively, compared with 116,500 and 88,400 amperes in the authors' paper.

With normal operation, the interphase transformer will delay build-up of reverse cathode current until it becomes saturated. This should require a matter of 20 to 40 electrical degrees, depending upon the interphase transformer design and the phase of arc back. However, current unbalance between the two wyes may result in an approach to saturation in the interphase transformer, and the pessimistic assumption of no delay in build-up of cathode current is sometimes realized.

REFERENCES

1. THE TESTING OF MERCURY-ARC RECTIFIERS, H. L. Kellogg, C. C. Herskind. *AIEE Transactions (Electrical Engineering)*, vol. 62, Dec. 1943, pp. 765–73.

2. See reference 5 of the paper.

J. K. Dillard and C. J. Baldwin, Jr.: We appreciate the comments of all the discussers. Their observations have contributed a great deal to a clearer understanding of the subject treated in our paper.

As Mr. Fraser points out, the use of larger capacity rectifier units, which implies lower reactance transformers, does result in higher arc-back currents. The reactance of the rectifier transformer has a substantial effect on the magnitude of fault current as evidenced by observing the effect of changing the per cent X_s on Figs. 11 and 12.

We regret that we were unaware of the model rectifier studies, using thyratron tubes, which Mr. Winograd described. In our search of the literature, the title of the paper he cited did not give us a clue that the problem had been studied on a model. Model studies are very useful for handling specific problems and installations. For general studies the analogue computer has an advantage because of the flexibility with which circuit parameters can be represented on a per-unit basis. In the section of the paper covering the analogue computer study some of the problems encountered in adapting a fixed arc drop, like that of a thyratron tube or germanium tube, to gathering general data are discussed. Using a fixed arc drop, all of the other circuit constants have to be represented as ratios of the fixed value, which limits the range of variables that can be considered Phase control studies can be handled on the analogue computer, although not quite as easily as on a model circuit.

A more realistic approach to handling the d-c feed into the fault has been suggested by Mr. Schmidt. Our representation of the d-c bus by a voltage fixed at 90 per cent of normal is defended on the basis of the difficulty which would be encountered in paralleling several sources having different internal impedances. If this difficulty is overcome through the separate representation of all the sources connected to the d-c bus, the complication of the study is increased tremendously and generality is lost. The chief limitations of selecting a fixed 90-per-cent voltage representation of the d-c bus are: 1. the results in many cases are too conservative; 2. the curves cannot be applied to a simple circuit where there are few parallel units or no counterelectromotive force to hold up the bus voltage.

Mr. Schmidt's other suggestions merit further consideration. We expect to extend our work to include the current in the normal anodes, as well as data on sympathetic arc backs. In making this study his comments relative to the assumption that the second normal anode enters conduction at the time when its potential becomes positive will be kept in mind.

Both our approximate method and the one referred to by Mr. Schmidt assume only two tubes conducting during the first half-cycle. The method we proposed has an advantage in making calculations in the field because no special equations or curves are needed. Also, this way of looking at arc-back faults may give a better physical picture to those of us who are accustomed to thinking in terms of faults on power systems.

Some Engineering Considerations in the Design of Telephone Systems to Serve Predominantly Rural Areas

T. J. McDONOUGH
ASSOCIATE MEMBER AIEE

WARNER T. SMITH
MEMBER AIEE

THE Rural Electrification Administration (REA) under an Act of Congress of October 1949 has the responsibility for making long-term loans for the construction of telephone systems in predominantly rural areas. Loans made under the Act are generally for a 35-year period at 2-per-cent interest.

Security for loans of this type is represented primarily by the ability of the system to provide good service at reasonable rates over the loan period. To insure that the systems for which loans are made represent adequate loan security, it is necessary to establish standards for the design and construction of such systems.

It is the purpose of this paper to discuss the more important phases of these standards. Information on the characteristics of the systems designed to date is provided as indicative of the problems involved and the need for developments to facilitate the expansion of rural telephony.

Average population density in the areas which are to be served is in some cases as low as 0.1 subscriber per square mile. The typical area is of the order of three subscribers per square mile. Seldom is there a populous area included within a system. The systems typically are located on the periphery of a populous area served by another organization which area is frequently the center of commercial and social interests of the subscribers to the rural system.

Basic Design Principles

In establishing the basic design principles to be used in the development of specific standards, the following are some of the more important considerations:

1. The objective of the REA program is to extend service to as many unserved rural areas as practicable.
2. Loan funds are available for a 35-year period at 2-per-cent interest.
3. The operating system provides the primary security for the loans.

Examination of these factors, along with those which apply to all rural telephone systems, leads to the formulation of the following basic principles to apply to all design activities:

1. The designs will be based on achieving the minimum over-all costs of rendering service, considering all factors.
2. The designs will provide for the rendition of service of good quality, as is essential to perpetuate an operating system as security for the loan.

In the application of the principle of lowest annual cost of designs, full advantage is taken of the relatively low annual cost of REA loan funds to defray other higher annual costs associated with operation, maintenance, etc. One example of the effect of this principle on designs is the preponderance of dial switchboards in designs.

In the usual case, the substantially higher investment required for dial switchboards compared to manual switchboards is more than offset by the lower annual cost of dial operations.

The more important REA design standards or objectives relating to quality of service are:

1. Adequate transmission and signaling.
2. Multiparty line fill of eight parties.
3. Full or semiselective ringing.
4. Adequate trunk facilities to interconnecting exchanges.
5. Twenty-four hour a day service, preferably of the dial type.
6. Plant design to achieve the maximum reliability of service.

Details of Design Standards

Details of design standards and objectives to carry out these basic principles are provided in the succeeding paragraphs. To a large extent, the design standards represent adaptation of applicable designs developed and used by the industry.

Fundamental Plan

A basic long range plan is required before attempting to design a system to provide telephone service. In order to develop this fundamental plan, a commercial survey is necessary to establish the immediate and future telephone needs of the area. With the results of such a survey, a design can be prepared which will indicate the type and extent of telephone facilities necessary to satisfy the requirements for telephone service. The design considers not only the initial demand for telephone service, but the anticipated growth for periods of 5 and 10 years in the

Paper 54-103, recommended by the AIEE Wire Communications Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 27, 1953; made available for printing December 4, 1953.

T. J. McDonough and Warner T. Smith are with the Rural Electrification Administration, Washington, D. C.

The authors wish to acknowledge the work of many of their fellow employees of the Rural Electrification Administration who played a major role in the development of the design standards discussed. The invaluable assistance of the telephone industry in making available necessary information so that the standards could be expeditiously prepared is gratefully acknowledged.

future. Where there is existing plant, the amount to be retained in the proposed system depends on its condition and its adaptability for use in the fundamental plan.

Central Offices

The number and location of central offices are determined by a subscriber loop and trunk study which is an attempt to place in equilibrium the annual costs of central-office equipment, buildings, trunk circuits, and subscriber loops. These studies are affected to a very great extent by the results of the commercial survey, and on the estimated calling habits of the existing and potential subscribers. It has been found that generally in rural areas telephone service can be most economically provided through the use of unattended dial offices working into larger attended offices which are owned by some other company.

Extended area service (EAS), toll free or flat rate, between central offices is employed in many instances to provide service between exchanges which have a substantial community of interest with each other. This is desirable where toll charges would not offset traffic and commercial costs.

Nation-wide Toll Dialing

System fundamental plans generally include provision for participation in nation-wide toll-dialing plans. For this reason, in addition to providing adequate transmission, switching, and signaling facilities, all central offices ordered since March 1953 are wired or equipped for 2-letter 5-digit numbering. Close liaison with the American Telephone and Telegraph Company and the operating Bell and independent companies is maintained to assure co-ordination with changes in nation-wide toll-dialing plans.

Transmission and Signaling

Toll trunks which are part of the intertoll dialing network have a transmission design objective of 4 decibels (db) net loss between the end office and the toll center. Trunks which are not part of the intertoll dialing network have a transmission design objective of 12 db net loss between the originating and terminating offices.

Loop signaling usually is not specified for 2-way trunks. Since only a small proportion of the central offices have sufficient trunks in a group to justify 1-way trunks, it finds little use in the program. The majority of trunk circuits are 2-way types using simplex or composite signaling and are designed to permit E-M duplex-type signaling whether or not it is initially required.

Transmission losses on subscriber loops are determined by using the loop loss factor method. Effective transmission loss of 0 db at the central office has been established as a design objective, and a line resistance of 1,000 ohms (excluding the telephone set) has been established as the signaling standard. Prior to the introduction of the Western Electric Company's type-*500* telephone set, the exchange transmission and signaling limits were compatible. The type-*500* set now makes satisfactory transmission possible on lines as high as 2,000 ohms which is beyond the capabilities of standard central-office equipment for supervision. At the present time, long-line adapters are required by most manufacturers to extend the supervision limits beyond 1,000 ohms.

Where required, use is made of cable loading and relatively high conductivity conductors to attain transmission objectives. On trunk circuits, voice-frequency repeaters and the improved transmission characteristics of carrier circuits are utilized.

Traffic

Interoffice trunking requirements vary depending upon the type of traffic handled. The design objective for engineering of toll trunks is not more than three to five lost calls out of 100 busy-hour calls. Extended-area service trunks are engineered so that there will be from two to ten lost calls in 100 depending upon the cost of providing the facilities, which in turn is related to the distance between offices.

Intraoffice trunking requirements for dial switchboards vary depending upon the number of switching stages. All are now designed so that for 98.5 per cent of the busy-hour calls, dial tone will be delayed 3 seconds or less.

Traffic estimates are based partially on existing calling habits and partially on engineering judgment. More than 400 switchboards were analysed and were found to be engineered on the basis of between 1.3 and 2.5 originating unit calls per line. In general, the higher the average line fill the greater the traffic. During the next year, we hope to check the estimated traffic against that which actually exists.

Ringing Systems

The objective of full-selective or semiselective ringing for multiparty lines requires that one of two available types of ringing systems be utilized. These are:

Multifrequency Ringing

This type of ringing system achieves its selectivity by mechanical and electric tuning of the station ringers to one of five frequencies transmitted by the central-office ringing machine. The frequency band used ranges from 16 to $66^{2}/_{3}$ cycles. There are various combinations of frequencies utilized by the several types of systems. The ringing frequencies can be generated by static, rotating, vibrator, or electronic types of machines.

With this system, five stations bridged can be rung with full selectivity. Ten bridged ringers can be rung on a semiselective basis. Divided connection of ringers, of course, will ring up to ten parties on a fully selective basis. Divided connection of ringers may be impracticable, however, where noise induction is important. As discussed under the section on dial central-office equipment, both divided and bridged ringing of multiparty lines can be used on the same system.

Certain lines where noise induction is a problem can utilize bridged ringing, while the remainder of the system can utilize divided ringing. Reconnection of ringers from divided to bridged can be done at any time when conditions warrant.

In the last 2 years, most manufacturers of "multifrequency" ringers have brought out new designs of ringers of the high-impedance type (over 100,000 ohms at 1,000 cycles). This type of ringer is beneficial not only from the noise-induction standpoint, in the event of divided ringing, but in addition, the lower current requirements of these ringers improve the reliability of ringing over the longer loops.

Superimposed Ringing

This type of ringing involves a single ringing frequency. It achieves its selectivity through divided connection of ringers, mechanical bias, and the superimposition of direct voltage. The ringing machines for this type of system consist of static or rotary type generators, with approximately 45 volts of direct current superimposed to give polarity selection. A divided connection will provide full selectivity for two ringers. The use of plus and minus polarity selections provides for full selectivity of four divided ringers. Semiselective ringing can be provided for up to eight parties by introducing a second ringing code. The station-equipment ringer is isolated from ground with a polarized cold-cathode tube.

When the correct polarity is applied, the tube breaks down and permits the ringing voltage to be applied to the ringers. Since the ringers are isolated from ground in the idle condition, they have no effect on noise. There have been some difficulties in the past where high induced voltages to ground were encountered, which caused faulty operation of the tubes in the idle period. A new tube containing four elements is now available, which will withstand a 180-volt peak between line and ground, and is adequate for all but the most unusual situations.

SYSTEM PROTECTION

As a design objective, all practicable protection means are utilized throughout the system to safeguard life and property. Outdoor-type station protectors are used at subscribers' premises. The protectors consist of 7-ampere power-rated fuses and a carbon-block type of low-voltage arrester connected to ground. The ground connection is made to the best ground available at the premises, and then interconnected with the power-service ground to obviate differences in ground potentials. The carbon blocks break down at about 450 volts and prevent higher voltages reaching the subscriber set. The fuses are present to protect against power contacts with peak voltages of less than 3,000 volts. They should not, but unfortunately often do, operate under surges.

Joint pole crossings are preferred to mid-span crossings with power lines. This minimizes the possibility of contacts in the event of power-line failure and, where the power system is of the multigrounded neutral type, permits access to a reliable ground for connection of power contact protectors which are recommended at all crossings with such systems. The power contact protectors consist of gaps from each wire to a ground electrode which break down at about 3,000 volts. The breakdown of the gaps protect the station fuses from having to clear voltages higher than their rated value of 3,000 volts and permits power-system fault current to flow to ground, thus obtaining rapid deenergization of the power system.

At the junction of open wire and cable carbon-block arresters are used with the cable terminals. These arresters have slightly higher breakdown voltages than those of the station protectors but are well within the breakdown voltage of the cable pairs to sheath. The protectors are grounded to the cable sheath and mitigate against high voltages getting into the cable from the open-wire plant. No fused type of cable terminals are used.

At the central office main frames, heat coils and carbon blocks are used for protection of the central office equipment. A short section of 24- or 26-gauge cable is installed as the entrance cable to limit currents to safe values. Much care is exercised to obtain a good central office ground.

In areas of high lightning incidence, pole protection grounds are installed at every fourth pole to minimize pole damage. These consist of wires which run from the top of the pole to a point near the ground.

Protection for joint use situations includes the use of power contact protectors at regular intervals tied in with the multigrounded neutral of the power system. Static voltage drainage units are used whenever required by the length of joint use circuits and other conditions.

DIAL CENTRAL OFFICE EQUIPMENT

There are six manufacturers whose central office equipment meets the general specifications established by REA. Three companies manufacture switch equipment and three manufacture all-relay equipment. All have equipment in service with one or more REA borrowers.

The general specifications require that the central office equipment will operate on lines with 1,000 ohms loop resistance at 68 degrees Fahrenheit (excluding the telephone set) with 15,000 ohms leakage resistance and 5-microfarad bridged capacitance. Full selective ringing for eight parties with multifrequency ringing requires divided ringing, but in some installations due to inductive interference full selectivity may be impracticable. The specification therefore requires that a 2-ring code be provided so that on such lines the ringers may be bridged and semiselective ringing used without requiring changes in the equipment or central office wiring.

Means are provided for transmitting an alarm indication to an attended location automatically as well as indicating the alarm condition locally. It is possible to determine from any telephone whether the alarm is of major or minor importance.

The nominal switchboard voltage is 48 volts d-c with batteries furnished which are capable of maintaining operation during failure of the a-c supply. Charging is on a full float basis with a full-wave, self-regulating constant voltage dry disk or equivalent type of charger being required.

Stand-by ringing, interrupter, and tone equipment are required to reduce as much as practicable the possibility of the central office being out of service due to failure of common equipment.

Type of Connectors

Two types of connectors may be specified in community dial offices. Terminal per-line connectors are capable of selecting a line and then selecting a station on that line. The number of stations that can be served by a connector terminal is determined by the number of different ringing combinations which the connector is designed to handle. The vast majority of the connectors for the more than 400 switchboards analyzed are capable of handling up to ten parties per line. Some of the larger central offices operate on a combination terminal per-line and terminal per-station basis. Some terminal per-station operation was selected because, while more costly for party lines, it has the following advantages:

1. Each connector is simpler and less costly.
2. Efficiency in operation of outside plant is increased.
3. Intercept equipment is simpler and less costly.
4. Subscriber may move yet the same directory number is retained.

Busy Verification

Busy verification is a feature which permits a toll operator to override a busy line indication. This is desirable to determine if a line is out of order or if there is actually someone talking. On toll calls it may also be used for number checking. The method of verification in use at the toll center is usually specified but wherever possible verification is done using the prefix digit "0" since level "0" on incoming toll selectors is available. Suffix digit verification is not recommended because of difficulties encountered in using it with certain senderized offices. Its advantage of not requiring a selector level is negated if prefix digit "0" may be used.

Revertive Calls

All of the dial switchboards are equipped so that parties on the same line may call each other without the assistance of an operator. Certain manufacturers find it desirable to complete revertive calls through the use of a special switch. This method requires dialing a special code of one or more digits followed by the last digit of the calling and called subscribers' directory numbers. The other method is to initiate a revertive call just like any other call by dialing the called subscriber's directory number. With either method the calling subscriber must hang up after dialing to start the called subscriber's telephone ringing. When the revertive call switch is employed, the calling and called subscribers' ringers operate alternately so that when the calling subscriber's telephone stops ringing it is an indication that the called party has answered. When revertive calls are completed by directory number, the calling subscriber must wait a reasonable length of time before lifting the receiver so as to give the called subscriber an opportunity to answer because ringing will stop when either party answers. So that the called party will know that it is a revertive call and wait for the calling party to get back on the line, a distinctive tone is given which indicates that it is a revertive call. It is relatively easy for the subscribers to determine who the other parties are on their line since in terminal per-line offices only the last digit of their directory numbers differ and the majority of community dial offices have terminal per-line connectors.

Lockout

Lockout is another feature required in community dial offices. Since the offices are usually unattended, an "off-hook" signal on a subscriber line due to a misplaced handset, a line cross, or a ring side ground could tie up switching equipment indefinitely. Line circuits equipped for lockout are automatically disconnected from the switching equipment after some predetermined time and remain disconnected until the cause is removed. On revertive calls with 100-per-cent lockout the line is put on lockout when the called party answers, thus making the switching equipment available for other calls.

Intercept

It is sometimes desirable to intercept calls directed to certain subscriber numbers or trunk levels. Intercept by an operator of a party-line station is the most complete type of intercept but for terminal per-line operation it is quite costly and is therefore usually not employed in terminal per-line offices of less than 400 lines. Operator, and later customer toll dialing, will make it desirable to have all unassigned directory numbers on intercept but with existing methods this is not economical in terminal per-line offices.

Central Office Specifications

The general specifications cover the standard features discussed in the foregoing. Specific requirements for individual offices are compiled and shown on a standard form. The manufacturers have become familiar with the requirements of the general specifications and the pattern which is developing in the detail specification so that engineering of individual offices is expedited. Further, it has led to advance manufacturing of certain components which is indicative of further possibilities in standardizing central office equipment.

EXCHANGE OUTSIDE PLANT

The cable portion of plant utilizes plastic-sheathed plastic-insulated cable or lead-paper type. Long-span lashed-aerial cable construction is commonly used. Because of unresolved problems connected with cable "dancing" under certain wind conditions, it has been necessary to limit span lengths in some sections of the country. Buried and underground cable have only limited application because the cable plant requirements of the systems do not justify economically such construction with present methods.

Open-wire plant is basically of the long-span type (300 feet average in heavy-loading areas and up to 400 to 500 feet spans in medium and light-loading areas). The wire is strung in accordance with conductor manufacturers' recommendations with tensions low enough to insure adequate fatigue endurance. The maximum span lengths are limited so that the breaking strength of the conductors are not exceeded under the assumed storm loadings of the National Electrical Safety Code (NESC).

The pole line design is based on not exceeding the fiber strength of the poles under the NESC-assumed storm loadings. This margin of strength is equalled or exceeded throughout all segments of plant on a co-ordinated basis. Deadends are given particular care to assure that they will not be the weak link in the line.

Conductors of high-strength copper-clad and galvanized-steel types are used on the basis of conductivity and transmission requirements or for specific corrosion resistance properties in areas having unusual atmospheric corrosion conditions. Spiral reinforcing splints of the factory-preformed type are used at points of conductor support to provide reliable ties which will minimize conductor vibration difficulties.

The open wires are transposed to the *R-1* scheme (transposed every other pole) on long-span construction and to the *R-2* scheme (transposed every fourth pole) where short spans are used. Tandem-type transposition brackets are used without ties. In some areas with high winds, tieing may be necessary to avoid conductor and insulator abrasion.

Open-wire construction exceeding two full crossarms (10 circuits) is seldom employed as cable circuits have been found to be more economical in this range of circuit requirements. The majority of open-wire pole line mileage is of the 1-circuit type. This type of construction has been given particular attention and a design using a 2-pin crossarm, without braces, type of construction was developed as an alternate to bracket construction. The 2-pin crossarm used has a rectangular notch to bind against the ungained side of the pole and when tightly bolted provides adequate resistance to canting. Tests indicate that with a 250-pound load on one end of the arm no objectional cant is experienced. The arm is 18 inches long and provides 12-inch spacing. It provides better configuration than brackets and results in 12-inch greater ground clearance than the bracket type. The crossarms are usually fabricated from stock not suitable for use for the longer crossarms and are therefore inexpensive. Recent cost data indicate that "in place" costs are comparable to that of brackets with associated straps. The crossarms are used also as tandem brackets for transpositions by mounting on the side of the pole.

Joint use with electric supply circuits is employed wherever practicable and economical. In the usual case, the economy will depend on the modifications required in existing electric pole lines. Where employed, the separations required by the NESC and the additional separations for long-span construction recommended by the Joint Pole Practices of the Edison Electric Institute and Bell System[1] are observed.

In all construction, the provisions of the NESC and local laws and ordinances are observed. Ground clearances on private property are worked out with property owners but in no case is less than 8 feet of basic ground clearance provided.

INDUCTIVE CO-ORDINATION

Every effort has been made in establishing the design standards to minimize the susceptiveness of the telephone plant to noise. The open-wire plant design and central office equipment circuit balance have been given specific study from the noise standpoint. As a result of these measures, it is believed that for the average exposure with normal power system influence noise problems will not be prevalent if the telephone plant is properly maintained.

TRUNK PLANT

Trunk plant designs are either open wire or cable and trunk circuits are con-

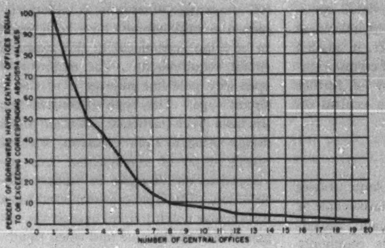


Fig. 1 (left). Number of central offices per system

Note: 115 borrowers and 438 central offices (applies to all figures)

Table I. Cost Distribution

Item	Approximate Per Cent for Rural Systems	Per Cent for Federal Communications Commission Class-A and -B Companies
Outside plant	72	39
Central office, including land and buildings	17	41
Station equipment	8	17
Other	3	3
	100	100

structed jointly with exchange plant where routing permits. For the shorter distances (10 miles or less), voice frequency circuits are generally favored. Phantom circuits are utilized to some extent in this range where power system exposures are favorable. Beyond this general mileage range, carrier circuits are considered and utilized when economical.

Trunk circuits use the exposed line transposition system. For carrier circuits, the transposition scheme used is dependent on the highest frequency of carrier that is contemplated. The United States Army line transposition scheme is now under study as a possible alternative to the exposed line system where no phantoms are involved.

In a few instances, point-to-point radio has been used to provide economical trunk circuits.

Quantitative System Design Data

Preliminary data that have been obtained from more than 100 systems designed to date are summarized now.

Number of Central Offices

Fig. 1 illustrates the distribution of central offices by number per system. All offices included in the study are of the dial type and are either on order or newly installed. It can be seen that 50 per cent of the systems have three or more central offices while only 20 per cent have six or more. One system has 20 central offices.

Subscriber Lines per Central Office

Fig. 2 shows the distribution of number of subscriber lines per central office. It can be seen from the figure that 60 per cent of the central offices have less than 80 lines but only 14 per cent have more than 200 lines. Only one central office has more than 625 lines.

Interoffice Trunks per Central Office

Fig. 3 shows the number of interoffice trunks per central office. It can be seen that 50 per cent of the central offices have between five and ten trunks with 42 per cent of the central offices having seven trunks or more. Less than 20 per cent have more than 12 trunks.

Subscribers' Data

1. The average system has about 1,100 subscribers.

2. The average central office has about 300 subscribers.

3. The average number of subscribers per mile of pole line is 3.5 ranging from less than one per mile in the Plains States to over 25 per mile in Louisiana.

4. The distribution of subscribers by grade of service is as follows:

1-party	18 per cent
2-party	8 per cent
4-party	19 per cent
multiparty	55 per cent

5. In a study of 11 system designs, it was shown that approximately 67 per cent of the subscribers are within 4 miles of the central office and less than 2.5 per cent of them are more than 12 miles from the central office. The introduction of the *500* series telephone and subscriber line carrier will extend loop limits operable from a single central office. Consequently it is expected that a smaller proportion of the subscribers will be within 4 miles of the central office and a greater proportion will be more than 12 miles from the office.

Pole Line Data

1. The average exchange pole line miles per system are 310 of which 290 are open wire and 20 miles are cable.

2. Over 50 per cent of the cable sheath mileage is 26 pair, with 22 per cent 51 pair, and 16 per cent less than 26 pair.

3. Open-wire pole line is broken down as follows:

2-wire	61 per cent
4-wire	14 per cent
6–10 wire	17 per cent
over 12-wire	8 per cent
	100 per cent

Extended Area Service Data

1. To date, approximately 100 systems have extended area service proposed or in operation with connecting companies.

2. The average distance between central offices is 10 miles.

Cost Distribution

Table I gives the figures obtained in a preliminary study of cost distribution for

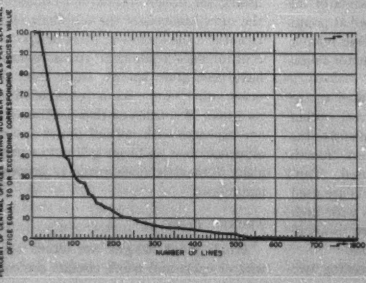


Fig. 2 (left). Number of subscriber lines per central office

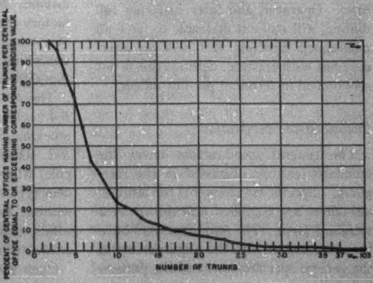


Fig. 3 (right). Number of interoffice trunks per central office

systems of REA borrowers compared to those of the larger operating companies in the telephone industry as a whole. It should be noted that data on class-*A* and -*B* companies include substantial toll and private branch exchange facilities which are not appreciable for rural systems.

Desirable Equipment Developments

Experience to date in the application of the design objectives to specific systems has emphasized certain segments of system design where equipment development would materially assist the job of extending rural telephone service. Many of these have been and are now under study by the industry. Some are in the immediate offing. A few of the more important development activities which it is believed would have major benefits are:

1. The type-*500* telephone set has minimized exchange transmission problems so that only the longest subscriber lines require special attention from this standpoint. However, standard central office equipment available to REA borrowers has not been modified to work with the resistance values obtained in such loops. Extension of central office loop limits to 1,500 ohms for reliable operation would provide for major economies in the outside plant design throughout the system.

2. Subscriber carrier has long been an obvious technique in achieving savings in outside plant and providing needed flexibility in satisfying demand for service in rural areas. Several manufacturers are now testing equipment of this type and the results of these tests will be of considerable interest to the industry. The equipment involved is low in first cost but as yet annual costs are unknown. Also, most of this equipment necessarily is wasteful of frequency spectrum space. If it fulfills the needs from cost standpoints, however, there will be time and opportunity to effect improvements in this field. If employed extensively, problems of frequency co-ordination involving facilities of separate organizations no doubt will arise. In these situations, it is hoped that the most economical solutions arrived at by engineering studies will be employed so that maximum economical use of all facilities results.

3. Plastic-insulated, plastic-sheathed cable appears to have a natural field of use in rural areas. Its lack of immediate sensitivity to moisture is particularly advantageous to the small operator with no skilled cable splicer available. This type of cable has, at present, less shielding than its counterpart in lead paper. As far as lightning protection is concerned, this is somewhat offset by higher dielectric strength between shield and pairs obtained with the plastic material. For noise induction, the low shielding characteristic is of some concern. The industry is actively studying this matter and improvements in shielding can be expected. Splicing techniques generally employed at present for this type of cable are not believed to be entirely satisfactory and improvement is needed. This is particularly true where splices are made to lead-paper cables which are moisture-sensitive. Some means of isolating the lead-paper conductors from plastic cable conductors moisturewise would be beneficial. Cable terminals generally have lead-paper stubs. When these are spliced into the plastic cable, similar difficulties with moisture may occur. Cable terminals with plastic-insulated stubs for use with plastic cables are an obvious need and are soon to be available. The same problem and need exists with case-mounted loading coils.

4. In cables of the smaller number of pairs, a relatively high percentage of installed cost is taken up with the messenger and lashing wire. Self-supporting cable would appear to offer economies in this field.

5. As mentioned under the section entitled "Ringing Systems," full selective ringing for as many as ten parties can be achieved at this time with divided connection of ringers. This type of connection poses problems because of noise considerations. Full selective ringing for ten parties without using the ground is desirable.

6. As pointed out in the section entitled "Quantitative System Design Data," there are a large number of central offices called for in system designs having less than 50 lines. For these central offices conventional designs result in relatively high investment per line. Specially designed central office equipment for this range of lines would be beneficial. Such equipment necessarily would sacrifice some features considered mandatory for larger offices. It is basic that substantial economy in common equipment will be required if over-all economy is to result. Pole-mounted equipment would appear to have advantages relating to savings in building and land costs.

7. In many rural areas the distances between central offices and between central offices and clusters of subscribers are large. Often no pole line exists in the area. In such cases point-to-point radio appears to have promise. Special equipment is needed for this purpose.

8. The Plains States of the country have areas where the population density is extremely low. There is a demand for service in these areas but with present techniques it is often unfeasible to provide service. Since power service usually is available in such areas, power-line carrier equipment appears to have possibilities if problems of initial and maintenance costs can be solved.

9. Rural areas of the nation have a demand for telephone service. The location of subscribers in relation to existing systems and to each other are materially different from that existing in urban and suburban areas. These differences suggest the need for a completely new look at system and equipment design in order to provide economical means of rendering service in such areas.

References

1. JOINT POLE PRACTICES FOR SUPPLY AND COMMUNICATIONS CIRCUITS, Joint Committee on Plant Co-ordination. *Report No. M-12*, Edison Electric Institute, Bell Telephone System, New York, N. Y., Oct. 1945.

2. AN IMPROVED CIRCUIT FOR THE TELEPHONE SET, A. F. Bennett *Bell System Technical Journal*, New York, N. Y., vol. XXXII, no. 3, May 1953.

3. THE DEVELOPMENT OF TELEPHONY IN THE UNITED STATES, A. B. Clark. *AIEE Transactions*, vol. 71, pt. I, Nov. 1952, pp. 348–64.

4. SOME ASPECTS OF JOINT USE OF WOOD POLE LINES WITH INCREASING LINE VOLTAGES, W. E. Bloecker, W. R. Bullard. *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 709–12.

5. GENERAL ASPECTS OF COMMUNITY DIAL OFFICE EQUIPMENT, Albert Burkett. *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1540–43.

◆

Discussion

H. R. Huntley (American Telephone and Telegraph Company, New York, N. Y.): The authors have given an excellent discussion of the rural problem and I emphasize their view that good objectives as outlined in their paper will ultimately result in better all round service. We in the Bell System have similar problems and in general are solving them in a similar manner. However, in a few cases, the details of engineering involved in carrying out these objectives as formulated in this paper differ in some respects from those of the Bell System. I believe that mention of these differences may be of interest to the Rural Electrification Administration and others interested in this paper.

In the Bell System, the selection of the number and location of central offices in an exchange area is generally a problem only in the larger cities. Studies which are made to determine the central office locations are known as "districting studies" in Bell System nomenclature.

In rural areas, the comparable problems are usually much simpler as the communities have ordinarily been established and the choice for each community is between a local central office or the extension of service from a larger near-by city. These choices are affected by many factors, including economics; but in Bell System such economic studies are not known as loop and trunk studies. This latter name has in the past been associated with outside plant cost balances covering loops and trunks in large cities. Such studies are no longer considered necessary because of the adoption in the Bell System of a new method for the design of loops. By virtue of the added efficiency of the new *500* type set, we have found that using the smallest gauge permitted by signaling and loading, all loops over about 18 kilofeet will give adequate transmission and eliminate the complicated study procedures of the past.

The matter of extended area with free service between central offices is a problem which has been given considerable thought in the Bell System. Recently new factors have come into the picture such as Automatic Message Accounting (AMA)[1] and Central Office AMA (CAMA)[1] which are expected to have a substantial influence on the cost balances between toll charges and traffic and commercial costs.

The intertoll switching network which is gradually being adopted as part of the overall nation-wide plan of the Bell System is also having considerable influence on the design objectives of trunks from the toll center to the so-called "end office." In the future, all intertoll trunks will be operated at via net loss (VNL) requiring that the toll link from the last toll center to the end office be operated as close to VNL plus 2 db as practicable and preferably not over 4 db. This objective may not be practicable at once in all cases, but with new, cheaper repeaters[2] and carrier systems[3] coming into the picture, should be attained within the next few years.

In the Bell System we are finding that 12 db is somewhat too high an objective for interoffice trunks, or, in fact, for any trunk not involved in the toll switching network. We have worked in the past toward an objective of 8 to 10 db and have recently reduced this objective to about 4 to 6 db for the immediate future (say 5 to 10 years). Also, we have found in the Bell System that a 2,000-ohm limiting loop, even if equipped with the new *500* set, would provide somewhat too poor transmission to be considered generally satisfactory. With the *500* set, we think we should not go beyond about 1,700 or 1,800 ohms. With regard to "busy verification," this feature may be very useful for operator toll dialing, but as far as we can see is not practicable for inward customer dialing sharing the same trunks, and would have to be eliminated when this type of operation is initiated. Likewise, "full intercept" is highly desirable, but may not be essential for operator dialing. However, it would be required after customer dialing is started and this would tend to discourage the use of terminal per line offices, looking forward to the future.

The question of liberality in provision of trunks is, of course, an economic one and will probably vary substantially between areas. Our view is that the objective liberality proposed in your paper; i.e., no more than 3 to 5 calls in 100 being delayed, would be very desirable particularly for customer dial operation, but may not be realized for some years because of economic reasons in particular cases.

In closing I wish to comment favorably on the very interesting statistical data which are provided in the latter part of the paper and also the valuable contribution in pointing up the needed developments which are listed in the last section.

REFERENCES

1. FUNDAMENTALS OF AUTOMATIC TELEPHONE MESSAGE ACCOUNTING, John Meszar. *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 255–69.

2. A NEGATIVE IMPEDANCE REPEATER, J. L. Merrill, Jr. *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1461–66.

3. TYPE-O CARRIER TELEPHONE, J. A. Coy, E. K. Van Tassel. *AIEE Transactions*, vol. 71, pt. I, 1952 (Jan. 1953 section), pp. 428–37.

Line Amplifiers for Community Television Systems

K. A. SIMONS
NONMEMBER AIEE

DON KIRK
NONMEMBER AIEE

H. J. ARBEITER
NONMEMBER AIEE

SINCE the first successful community television system was installed in May 1950 at Lansford, Pa., this infant industry had made a great advance. At present there are some 240 systems providing television reception to a potential audience of several hundred thousand people throughout the country.

This mushroom growth has brought with it the development of many new techniques and the need for solving a multitude of technical problems. Although some of the art is entirely new, the basic technique of moving television frequency signals efficiently from one place to another is closely allied with long-distance telephony.

The possibility of a community television system exists where a large group of people who cannot receive satisfactory signals directly are located within a few miles of a site where satisfactory reception is possible. The most obvious case is a town shadowed by a mountain, as shown in Fig. 1. The system operator installs high-gain antennas and sensitive receiving equipment on top of the mountain where the signals are relatively strong, and distributes them by coaxial cable to the homes in the town. Since these signals must be delivered to the customers on the standard television channels between 54 and 216 megacycles (mc), present practice is to select channels most suitable for the system, and to convert the frequencies of the received signals wherever necessary to fit this pattern. Most of the systems designed by Jerrold Electronics Corporation have used channels 2, 4, and 5 (or 6) because of the lower cable loss on these lower frequencies. A 7-channel system, with nonadjacent channels, is possible using channels 2, 4, 6, 7, 9, 11, and 13. These two cases will be considered in this article. Although operation of adjacent channels is possible, it has not been widely used because of the difficulties of preventing interference between channels, and it will not be considered.

This article investigates one of the major problems of community system design, the choice of suitable amplifiers for relaying signals along the main trunk line. Three important possibilities are considered:

1. Separating channels frequencywise at each repeater point and amplifying each with a separate single-channel cascade amplifier.
2. Separating channels into two groups and amplifying those between 2 and 6 with one broad-band cascade amplifier and those between 7 and 13 with another.
3. Using a single distributed amplifier for all channels.

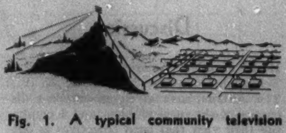


Fig. 1. A typical community television system

Paper 54-101, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 27, 1953; made available for printing December 1, 1953.

K. A. SIMONS, DON KIRK, and H. J. ARBEITER are with the Jerrold Electronics Corporation, Philadelphia, Pa.

Table I. Calculated Maximum Output Voltage of Band-Pass Amplifiers

Assuming Tube with Output Capacity of 5 Micromicrofarads, Output Current of 4.5 Milliamperes Rms, Optimum Output Transformer, and a 72-Ohm Line

Amplifier	Channels	Bandwidth, Mc	Output Voltage	Relative Level, Db
Single-channel cascade	any one	6	3.8	0
Low broad cascade	2-6	34	1.6	−7.6
High broad cascade	7-13	42	1.4	−8.5
Distributed: 6 tubes per stage	2-13	162	2.2	−4.9

Table II. Minimum Input Voltage, Maximum Gain, and Spacing

Assuming 30-Db Ratio of Picture White to Noise and the Use of *RG11/U* Cable

Amplifier	Channel No.	Assumed Noise Figure, Db	Minimum Input, Microvolts	Maximum Output, Volts	Maximum Gain, Db	Maximum Spacing, Feet
Single channel	6	6	500	3.8	77.5	3,950
Single channel	13	10	800	3.8	73.5	2,300
Cascade broad-band	2-6	8	600	1.6	69	3,500
Cascade broad-band	7-13	10	800	1.4	65	2,030
Distributed, 6 tubes per stage	2-6	6	500	2.2	73	3,700
Distributed, 6 tubes per stage	2-13	10	800	2.2	69	2,160

A number of the factors affecting this choice are developed, basing the calculations on reasonable but somewhat simplified assumptions. As is often the case with investigations of this type, this one is a little late. Several years of experience have established the single-channel amplifier as the best choice in a large majority of the existing systems. However, a theoretical investigation will serve to show some of the reasons for this choice and will indicate conditions under which it might not be valid.

Factors Limiting Allowable Spacing Between Amplifiers

Maximum spacing between the line amplifiers is desirable because it minimizes the number of points that must be made accessible for maintenance and the number of power-line connections. Maximum spacing is obtained by reducing cable loss to a minimum and utilizing the maximum possible amplifier gain.

CABLE LOSS

Cable loss is an inverse function of cost, more expensive cables with more copper giving the lowest loss. In practice, the cable used has been selected to give the best compromise between cable cost, equipment cost (which decreases as better cables are used) and maintenance. In a majority of installations *RG11/U* cable is used for the trunk line. Over difficult terrain where accessibility is a major problem more expensive cables such as *RG35/U* have been used.

In the attempt to increase the spacing between amplifiers, amplifier gain cannot be increased indefinitely. A limit is set by the necessity that the signal level at the input to each succeeding amplifier shall safely override the noise of that amplifier. Maximum spacing is obtained by making the amplifier output level as high as circumstances permit and obtaining the best possible noise figure in the input stage.

OUTPUT VOLTAGE

In a single-channel amplifier handling a picture carrier and its associated sound carrier, the output voltage into the line can be increased to the point where the video signal cross-modulates the sound. This is manifested as 60-cycle amplitude modulation of the sound carrier and it produces annoying buzz in receivers which are mistuned or which have inadequate limiters. In broad-band amplifiers handling several channels the limit on output voltage is set by cross-modulation between picture signals. This becomes serious at somewhat lower levels than those giving sound buzz and it results in moving bar patterns (windshield wiper effects) on the received pictures.

It is present practice in amplifiers for community systems to use small receiving tubes, chiefly *6AK5* or *6CB6*, throughout. The output stage is a single tube operated class *A*. With these limitations the output voltage into the cable depends primarily on the bandwidth requirements and on the efficiency of the output coupling network.

Table I shows the calculated maximum output voltage into a 72-ohm cable obtainable from the commonly used arrangements, assuming about 5-per-cent intermodulation and a tube with approximately the characteristics of a *6AK5* or *6CB6*. See Appendix I for method of calculation. The figures are optimistic in that they assume a lossless output circuit of optimum design, a condition which is easier to approach in broad-band circuits. Also, they disregard the reduction in allowable level per channel when a number of channels are present in the output of a broad-band amplifier. This condition does not apply to single-channel amplifiers.

INPUT VOLTAGE

The lowest input voltage which may be permitted into a line amplifier depends on the noise figure of the amplifier and on the lowest permissible signal-to-noise ratio. The minimum input-voltage figures in Table II have been calculated from the noise figures shown, based on a noise bandwidth of 5 mc and a ratio of 30 decibels (db) between noise and white level in the picture signal. By dividing these input voltages into the maximum output for each type of amplifier, the maximum possible gain is found. The cable loss between amplifiers must equal the amplifier gain, so the gain figures have been used to calculate maximum amplifier spacing. The use of *RG11/U* cable is assumed.

Minimizing Gain Variations Caused by Tubes

The operation of a repeater system depends on a critical balance between two factors, cable loss and amplifier gain. Variations from normal in either require means to readjust levels, and sufficient excess gain to allow system operation when cable loss is at its maximum and amplifier gain at its minimum. Even worse, these variations subtract from the allowable loss between amplifiers. Assume, for example, amplifiers with a spread of 60 db between minimum permissible input and maximum output. If signal level, amplifier gain, and cable loss were all constant, these amplifiers could be spaced 60 db apart on the cable. A variation in any one of these three means that the amplifier spacing must be reduced accordingly. A signal variation of 10 db would require reducing the spacing to 50 db of cable, with corresponding gain reduction in the amplifiers, to prevent the signal from varying beyond the limits set by the input and output circuits.

The worst source of fluctuations in a system, except for fading of received signals, is the change in tube gain caused by

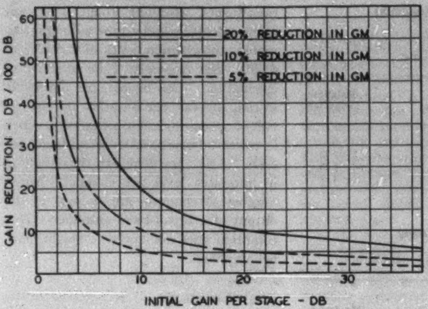


Fig. 2. Gain reduction caused by reduced mutual conductance

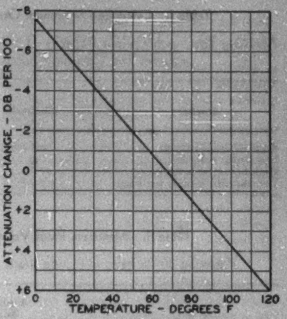


Fig. 4. Calculated change in cable attenuation with temperature

variations in mutual conductance. This results from supply voltage changes and from aging, and can be minimized by regulating supplies and by periodic replacement of failing tubes. Even with the best design and maintenance, it remains a major source of trouble.

A striking fact, and one of great importance in system design, is that tube variations cause less trouble in high-gain stages than in low-gain stages. This results from the fact that a given per-cent change in mutual conductance changes the gain of a stage by the same number of db, regardless of the stage gain. In Fig. 2 this effect is demonstrated by plotting amplifier gain reduction, measured in db per 100 db of initial gain, against the gain per stage. Note that amplifiers with gains of 20 db per stage or more are quite insensitive to variations in mutual conductance, whereas those with gains below 10 db are subject to much greater trouble from this source.

Fig. 3 shows an experimental check on this. The gain of each of two amplifiers was measured, and the line voltage was varied over the range from 125 down to 100 volts, using unregulated power supplies. The broad-band amplifier, which had a gain of about 5 db per stage, suffered a reduction of more than 17 db per 100 db of initial gain, while the single-channel amplifier, with a gain of about 15 db per stage, had its gain reduced less than 5 db per 100 for the same reduction in line voltage. The amplifier with the higher stage gain would be similarly insensitive to any other factors tending to change the individual tube characteristics, such as aging and tube replacement.

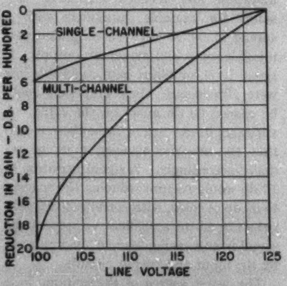


Fig. 3. Measured gain reduction of two amplifiers

The amplifier types being considered in this article can be compared in regard to their susceptibility to tube changes by calculating the stage gains to be expected. The results of these calculations are shown in Table III. They assume a gain–bandwidth product of 200 mc for the cascade amplifiers, and of 100 mc multiplied by the number of tubes per stage in the distributed amplifier. See Appendix II for method of calculation.

Control of Cumulative Effects on Long Systems

Level Variations

It is an axiom of repeater system design that amplifier performance requirements become increasingly difficult to meet as the length of the system is increased. One of the hardest requirements to meet is that of gain constancy. It is possible, by careful design and maintenance, with the use of regulated supplies, to hold the gain of single-channel amplifiers within a tolerance of perhaps ±3 db per 100 db of gain. This means that the level at the end of a 500-db system (about 5 miles of *RG11/U*) would vary over a 30-db range as a result of this alone. A further source of variation which can give trouble on long systems is the change of cable attenuation with temperature. The chief cause of this change is simply the change in the conductivity of the copper conductors.

Fig. 4 shows the calculated relation between cable attenuation change and temperature, based on the change of copper conductivity. See Appendix III. This curve indicates that normal day-night temperature changes of the order of 20 degrees Fahrenheit are sufficient to account for attenuation changes of 2 db per 100, which is enough to cause trouble on a long system.

The most effective tool for combating these gain and loss variations is the use of automatic gain control. This operates to hold the output level of the controlled amplifier almost constant for wide variations in input. By using automatic gain control at frequent intervals along the system (present practice in Jerrold-engineered systems is to control every third amplifier), gain variations are canceled out before they can accumulate to give trouble. As long as the amplifiers between controlled points are conservatively spaced so that they can tolerate their own gain variations and those caused by intervening cable, such a system can be extended indefinitely without trouble from gain variations.

While automatic gain control is relatively easy to apply to a single-channel amplifier, it is difficult to find a satisfac-

Table III. Calculated Gain per Stage and Gain Variability

Amplifier	Gain per Stage, Db	Gain Reduction for 10-Per-Cent Reduction in *Gm*, Db per 100
Single channel	30.4	3.3
Cascade broad-band		
channels 2 to 6	15.4	6.5
channels 7 to 13	13.6	7.4
Distributed amperes, 6 tubes per stage, channels 2 to 13	11.4	8.8

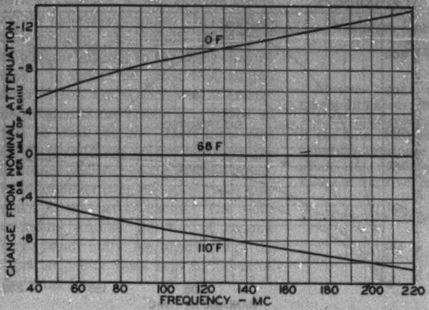


Fig. 5. Calculated change in attenuation with temperature per mile of *RG 11/U*

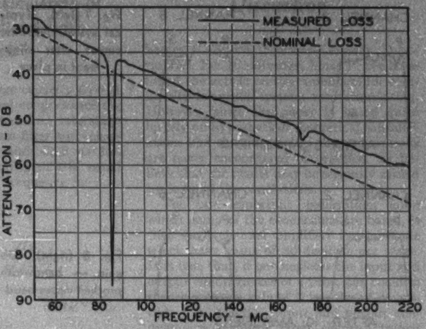


Fig. 6. Loss versus frequency for 2,000 feet of defective *RG 11/U* cable

tory way of controlling a multichannel unit. The relative levels of the various channels change, so controls based on any one channel are apt to allow noise or overload on other channels. Controls which respond to total voltage due to all channels are apt to be taken over by a strong channel, driving the others "into the snow" (i.e., reducing signal level below noise).

Perhaps the most workable solution is to break the system into separate channels at each point where automatic gain control is used, but in cases where this is practical, it is generally more satisfactory to use single-channel amplifiers throughout the system.

Frequency Response Variations

In multichannel amplifiers of the type used for community television the interstage networks are broadly tuned to cover wide bands. For this reason there is little difficulty in keeping the response over any one channel sufficiently flat to prevent frequency discrimination in the picture. The problem with multichannel amplifiers is to keep the wide-band response flat enough to prevent variations in the relative levels of the various channels. If this is not done, troubles caused by gain variation are exaggerated and automatic gain control is made impossible. Since cable attenuation varies with frequency, it is necessary to correct the amplifier response for this, to keep the total response uniform. This is commonly done with cascade amplifiers by aligning them with the correctly tilted response, and with distributed amplifiers by using a correction network which inserts additional loss at the lower frequencies as required to eliminate the tilt. Both schemes depend on agreement between the cable tilt in one direction and amplifier, or network, tilt in the other.

Any variation in cable attenuation slope will upset this relation. Unfortunately, the attenuation curve of various samples of cable varies considerably; and the attenuation curve of any one piece of cable is affected by temperature. Fig. 5 shows the variation from nominal attenuation over the frequency range of interest for temperature extremes to be expected. This is based on the curve shown in Fig. 4. Since the effect of temperature depends on the initial loss, it is greater at high frequencies where the initial loss is high than at low frequencies where it is lower. Thus, wide temperature changes introduce a tilt in the frequency response which is difficult to take into account in system design.

In single-channel amplifiers the level variations between channels, caused by cable characteristics, are corrected by the automatic gain controls; and a much less stringent requirement, that the response be maintained flat on each individual channel, controls the design of the system. Cable response does not usually vary enough over any one channel to give trouble; however, defects in cable manufacture occasionally cause severe irregularities.

Fig. 6 shows the frequency response of a reel of cable rejected as unsuitable for community work, for a reason which is apparent from inspection of the curve. An unfortunate periodic effect in the manufacture of this particular piece of cable resulted in a response some 50 db below normal, in the middle of a channel. Fortunately, this kind of thing does not often happen, but it is enough of a hazard to make it highly desirable to check the frequency response of each piece of cable to be used in a community system before it is installed.

Generally, it has been found possible to hold the frequency response curves of individual channel amplifiers sufficiently flat in production to prevent cumulative response troubles. In cases where extremely long runs are necessary (more than 10 miles is considered "extremely long" in the present state of the art) or where superlative performance is desired on shorter systems, it has been found quite practical to align each amplifier in place in the system, using a sweep-frequency generator at the antenna site, thus canceling out any irregularities and obtaining relatively flat curves throughout the system.

Fig. 7 shows typical response curves obtained on a system aligned in this fashion. It may be observed that there is almost no cumulative deterioration of the system response as additional repeaters are added.

Minimizing the Number of Tubes in a System

If the attempt is made to design a system for minimum cost, minimizing the number of tubes of a given type would appear to be a reasonable approach. The theory involved in this approach is quite simple, and the results clarify important factors affecting the choice of amplifiers.

Requiring maximum channels with minimum tubes of the individual amplifier stage has the effect of creating a figure of merit for the stage which is the product of the stage gain, in db, by its bandwidth in mc. This may be called the "db gain-bandwidth product" to distinguish it from the more familiar voltage gain-bandwidth product. For a given tube type and coupling network and certain

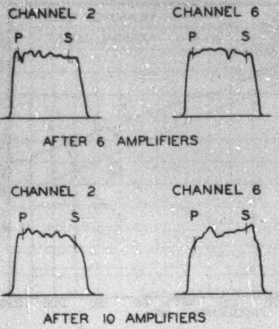


Fig. 7 (left). Typical system response curves after alignment in place

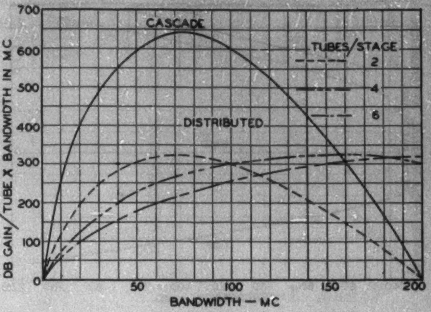


Fig. 8 (right). Db gain — bandwidth product versus bandwidth

simplifying assumptions, the latter product is a constant. Since db gain is related to the logarithm of voltage gain, this means that the db gain–bandwidth product, is a variable depending on bandwidth.

Fig. 8 shows this product plotted against bandwidth for single-tube cascade amplifier stages, assuming a voltage gain–bandwidth product of 200 mc, and for distributed amplifiers with various numbers of tubes per stage, assuming a voltage gain–bandwidth product of 100 mc multiplied by the number of tubes per stage. The product is divided by the number of tubes in a stage in each case to permit direct comparison. Each curve reaches its maximum at the bandwidth where the gain per stage has the optimum value of 8.7 db, or 1 napier.

With these assumptions it is apparent that the cascade amplifier is more effective for bandwidths less than about 160 mc, the distributed circuit being more effective at greater bandwidths and providing the only possibility of gain for bandwidths of more than 200 mc.

Conclusions drawn from these curves cannot be applied to the community television problem without reservations. Obviously, added bandwidth in an amplifier is desirable only when it adds a useful channel. An amplifier would require a bandwidth of 34 mc to pass channels 2, 4, and 6; or 42 mc to pass 7, 9, 11, and 13. Adding bandwidth in either case to reach 73 mc, the optimum according to the curve, would serve only to decrease the gain per stage for the desired channels, with no compensating advantage.

A more realistic approach to the problem of system economy involves finding the number of tubes required per mile of cable for each amplifier type and channel arrangement. Table IV shows the results of these calculations, assuming amplifier types with the gain characteristics shown in Fig. 8. The amplifier gain has been assumed to be just great enough in each case to overcome cable loss at the highest frequency to be amplified.

Table IV. Tubes Required per Mile of *RG 11/U*

	Tubes per Mile	Tubes per Channel-Mile
3-channel system using channels 2, 4, 6		
Cascade 54-88 mc amplifier	6.9	2.3
3 single-channel amplifiers	9.3	3.1
Distributed 54-88 mc amplifier (2 tubes per stage)	13.7	4.6
7-channel system using channels 2, 4, 6, 7, 9, 11, 13		
Cascade 54-88 mc amplifier plus 174-216 mc amplifier	19.8	2.8
7 single-channel amplifiers	31.3	4.5
Distributed amplifiers (54-216 mc, 4 tubes per stage)	54.0	7.7

Table V. Summary of Performance Characteristics

Numerals Indicate Order of Preference in Regard to Each Characteristic

Performance Characteristic	Single-Channel Amplifier	Cascade Broad-Band	Distributed Broad-Band
Output voltage	1	3	2
Input voltage	1	2	1
Spacing along cable	1	3	2
Gain variation	1	2	3
Use of automatic gain control	1	2	2
Control of frequency response	1	2	3
Economy of tubes	2	1	3

The cascade broad-band amplifiers show up well on this basis because their bandwidths fall near the optimum value, and each band is well filled with useful channels. The distributed amplifier suffers from a lower voltage gain–bandwidth product, per tube, and from the fact, in the case of the seven-channel system, that it amplifies a wide band between 88 and 174 mc where there are no channels.

Dividing the figures in Table IV for number of tubes per mile by the number of channels in each case gives an informative result, the number of tubes per channel-mile. This figure allows a comparison between systems having different numbers of channels. It will be noted that the 7-channel system shows greater tube requirements on this basis, reflecting the inefficiency of distributing signals on the higher frequencies where cable loss is increased.

Conclusions

Table V summarizes the various factors discussed in this article, as they affect the choice between the three basic types of amplifier suitable for line amplifiers in community television systems. Overwhelming evidence favors the single-channel amplifier. Using the same tube types, and making other reasonable assumptions, it has been shown that this type is capable of greater output voltage and that it requires no greater input voltage than other types, and thus it allows greater spacing between amplifiers with a given cable. Since it has higher stage gain it is less susceptible to gain variations caused by tubes. Automatic gain control may be used much more readily with single-channel amplifiers, and they allow simple correction of frequency response variations affecting relative levels between channels, as well as those affecting the individual channel. In regard to tube economy, the cascade broad-band amplifier is superior to the single-channel type. It should be noted that all the other factors are of increasingly greater importance as the system is extended, so that on small systems, where cable runs are short and there are few repeaters, the increased economy of the cascade broad-band amplifier over the other types may make it the better choice. The same reasoning justifies the use of cascade broad-band units for distributing the signal at points where taps are taken off the trunk line.

Appendix I. Calculation of Output Voltage of Band-Pass Amplifiers

The following assumptions are made:

C_p (output capacity of tube) = 5 micromicrofarads

R_L (load resistance) = 72 ohms

E_L = load voltage (rms)

I_p (maximum variation in plate current) = 4.5 milliamperes rms. (This figure was obtained graphically from tube characteristic curves for *6AK5* and *6CB6*, and represents plate current swing at 5-per-cent intermodulation point.)

E_p = plate voltage variation (rms)

The output transformer is assumed to consist of a band-pass filter with characteristic resistance R_0 and C_p as a full shunt arm followed by a loss less transformer matching R_0 to R_L. BW (bandwidth for flat response) $= 0.9f_c$ (cutoff frequency of filter).

For a Single-Tube Output Stage

$$R_0 = \frac{1}{\pi f_c C_p} = \frac{0.9}{\pi BWC_p}$$

$$E_p = I_p R_0 = \frac{0.9 I_p}{\pi BWC_p}$$

$$E_L = E_p\sqrt{\frac{R_L}{R_0}} = I_p R_0\sqrt{\frac{R_L}{R_0}} = I_p\sqrt{R_L R_0}$$

$$= (4.5)10^{-3}\sqrt{\frac{(72).0.9}{\pi(BW)5(10^{-12})}}$$

$$= \frac{9.2}{\sqrt{BW_{mc}}}\ \text{volts}$$

For the Output Stage of a Distributed Amplifier

The output current of each tube splits two ways, one half flowing in the output load. For the same bandwidth the band-pass filter will have the same constants as those previously mentioned. With n tubes the total plate current swing will be $(n/2)I_p$. Thus the load voltage will be

$$\left(\frac{n}{2}\right)\frac{9.2}{\sqrt{BW_{mc}}} = \frac{4.6n}{\sqrt{BW_{mc}}}$$

Appendix II. Calculation of Voltage Gain–Bandwidth Product

Assume: total interstage capacity $C_t = 14$ micromicrofarads; mutual conductance $G_m = 5{,}000$ micromhos; 4-terminal interstage in the form of a band-pass filter with characteristic resistance R_0 and $C_t/2$ as full shunt arm, response bandwidth BW equal to 0.9 of the filter cutoff frequency f_c.

For Cascade Stage

$$\text{Gain per stage} = G_M R_0 = \frac{G_M}{\pi f_c C_T/2}$$

$$= \frac{0.9 G_M}{\pi BWC_T/2}$$

$$\text{Gain}\times\text{bandwidth} = \frac{0.9G_M}{\pi C_T/2}$$

$$= \frac{0.9\times 5{,}000\times 10^{-6}}{\pi\times 7\times 10^{-12}} \approx 200\ \text{mc}$$

For Stage in a Distributed Amplifier

The plate current of each tube splits two ways, so one half flows into load resistance of filter. With n tubes in a stage the gain is $n/2\times$cascade stage gain.

Thus gain×bandwidth $= (n/2)\times 200$ mc $= 100\ n$ mc.

Appendix III. Calculation of Change in Cable Attenuation with Temperature

Neglecting dielectric losses,[1] the cable attenuation A is proportional to the total radio-frequency resistance of conductors R_{ac}.

Since the radio-frequency resistance of a cylindrical or tubular conductor is proportional to the square root of the resistivity[2] for fixed geometry

$$A \propto R_{ac},\ R_{ac} \propto \sqrt{R_{dc}},\ \text{so}\ A \propto \sqrt{R_{dc}}$$

but[3]

$$R_{dc} = R_{20}[1+0.00393(T-20)]$$

where

$R_{20} = dc$ resistance at 20 degrees centigrade

T = temperature degrees centigrade

$$\frac{A}{A_{20}} = \sqrt{\frac{R_{dc}}{R_{20}}} = \sqrt{[1+0.00393(T-20)]}$$

Attenuation change in db per 100 db of

$$\text{initial attenuation} = 100\left[\frac{A}{A_{20}}-1\right]$$

$$= 100\left[\sqrt{1+0.00393(T-20)}-1\right]$$

References

1. Reference Data for Radio Engineers (book). Federal Telephone and Radio Corporation, New York, N. Y., 3rd ed., 1949, p. 319.

2. *Ibid.*, p. 36.

3. *Ibid.*, p. 41.

Discussion

Lester C. Smith (Spencer-Kennedy Laboratories, Inc., Cambridge, Mass.): We have found the article on line amplifiers for community television systems very stimulating. However, since the paper is based quite largely on hypothetical theoretical considerations and since several of the assumptions are never encountered in practice, we believe that the reader may be seriously misled by some of the conclusions.

The one important assumption, upon which many of the conclusions are based, is that single-channel amplifiers can be made which have gains of 30 decibels per stage. It is true that, if one neglects the input and output conductances of a tube, one could in theory secure a gain of this order of magnitude. In practice, input and output conductances resulting from electron transit time and electrode lead inductances are quite serious at these frequencies. In addition, stable single-channel amplifiers, which maintain their adjustment for a reasonable period, must have additional capacitance at the tube electrodes, with the result that no single-channel amplifiers are being produced which have stage gains exceeding approximately 15 decibels; the better quality units have gains of approximately 10 decibels, a fact which may be verified by examining the bulletins of manufacturers of this type of amplifier. Because of the special characteristics of broad-band distributed or "chain" amplifiers, excellent stability is obtained without added capacitance at the tube electrodes; and amplifiers which have gains of 10 decibels per stage are found to be thoroughly practical and dependable.

Since the two types of amplifiers now being produced have approximately the same gain per stage, the gain stability of the single-channel type is no better than that of the distributed type. The figures in Table III and the curves in Figure 3 are therefore misleading. Likewise, the data in Table IV are not true in practice since, for example, the tubes per channel-mile for a 7-channel system would be approximately 18 for the single-channel amplifier and approximately 15.5 for the distributed or chain amplifier.

The output levels which the authors deduced and reported in Table I are also misleading because, even in the case of the single-channel amplifier, cross-modulation between picture and sound carrier would become intolerable in a system of moderate size unless the signal were at much lower levels than that assumed in the article. In practice, operators of single-channel systems have found it necessary to operate with signals 20 db lower than are given in this article. (We should like to add that the maximum output per channel for a distributed amplifier, when carrying seven channels, is also approximately 20 db less than the figure the authors used.)

Another important factor neglected is the build-up of tube noise in a cascade of amplifiers such as is used in a community television system. Basic theoretical work, indicates that optimum performance is obtained when amplifier gain is quite small, so that the signal level at the input to each amplifier is far above the noise level. The optimum input signal level depends upon the length of the system but, in any case, it is very many db higher than that assumed by the authors.

As a consequence of the fact that allowable output levels are approximately 20 db lower than quoted in this paper, and desirable input levels are 10 to 30 db higher than the figures in the paper, the optimum gain for an amplifier, whether single-channel or distributed, is in the 20- to 40-db region and not in the 60- to 80-db region given in Table II. Since the paper was largely theoretical, I will not mention the many practical advantages which we have found the distributed or chain amplifier to possess.

K. A. Simons, Don Kirk, and **H. J. Arbeiter:** The authors appreciate Mr. Smith's interest in the paper. Perhaps the most adequate response to his criticisms of the theoretical approach used in the paper is the inescapable fact that our conclusions agree with experience gained in the installation of nearly 240 successful community antenna systems. We admittedly simplified our initial assumptions in the attempt to establish goals for future design, rather than merely to catalog the limitations of present equipment. This approach has proved to be both informative and useful. Since the article was written, we have begun producing amplifiers (for our new 5-channel system) capable of nearly a 20-db gain per stage and output voltage in excess of 1.5 volts. There is no reason to believe that progress will stop here.

The single-channel curve in Fig. 3 was measured on an early sample of one of these amplifiers. It is interesting to note that these curves, labeled misleading by Mr. Smith are plots of careful measurements on production model amplifiers! In fact, recent measurements of the variation of gain with line voltage on a commercial distributed amplifier indicate that its performance agrees almost exactly with the multichannel curve on Fig. 3.

Actually none of this discussion has touched on what we feel to be the most vital factors affecting the choice of amplifier types. As indicated in the conclusion of our paper, the problem naturally divides into two major cases as follows:

1. Long systems (of more than 5 miles or so), in which the difficulty of maintaining constant level and flat frequency response after many repeaters overrides all other considerations. Because a single-channel amplifier allows simple automatic gain control for constant level, and simple alignment for flat frequency response, we are firmly convinced that it is the unqualified choice for this application.
2. Short systems, in which there are few repeaters. In this case cost is the controlling factor, and the best system is the one with the fewest tubes per channel-mile. From this standpoint both theory and practical tests indicate the cascade broad-band amplifier to be at least twice as good as the other types.

Although we are fully conscious of the outstanding advantages of the distributed amplifier, particularly where extreme bandwidths are required, we are not convinced that it has application in community antenna systems.

Design Considerations of the Half-Wave Bridge Magnetic Amplifier

C. W. LUFCY
ASSOCIATE MEMBER AIEE

H. H. WOODSON
ASSOCIATE MEMBER AIEE

AN improved magnetic servo amplifier utilizing half-wave bridges in cascade has been extensively used in instrument-type servos designed at the U.S. Naval Ordnance Laboratory. The speed of response inherent in a magnetic amplifier of this type has made possible the application of conventional compensation and stabilization techniques.[1] The result is dynamic servo performance characteristics far exceeding those of previous magnetic-amplifier servo systems employing magnetic amplifiers using full-wave circuitry.

In a previous article[2] the basic half-wave circuit was described along with information to permit construction of a 60-cycle and 400-cycle servo amplifier to demonstrate its advantages. The specific amplifiers therein described did not represent optimum design nor were considerations to achieve optimum design discussed. It is the purpose of this paper to describe the operation in such a manner that optimum design and performance of this circuit can be approached.

Description of Half-Wave Bridge Magnetic-Amplifier Circuit

The half-wave bridge circuit is shown in Fig. 1. The windings N_1 and N_3 are the power windings on one reactor while N_2 and N_4 are power windings on the second reactor. These two reactors make up the bridge. The windings N_{r_1} and N_{r_2} are the reference, or reset, windings on each reactor. Their purpose is to reset a predetermined flux level during the control half-cycle when conduction is prohibited in the power windings by rectifiers R_{X_1} to R_{X_4}. Control of the bridge is effected by a signal on the control windings N_{c_1} and N_{c_2}. In Fig. 1 all windings having an odd subscript are common to one reactor while all windings with an even subscript are common to the other.

Briefly, the operation of the bridge is as follows. Under zero signal conditions the bridge stays balanced, because during the half-cycle (the reset half-cycle) when rectifiers R_{X_5} and R_{X_6} conduct, the flux change from saturation in each reactor, produced by voltage across windings N_{r_1} and N_{r_2}, is the same. As a result, during the next half-cycle (the operating half-cycle) when rectifiers R_{X_1} to R_{X_4} conduct both cores reach saturation at the same time, maintaining zero load current. With the application of a signal across the differentially connected control windings N_{c_1} and N_{c_2}, a differential flux is established between the cores in the reset half-cycle. This difference in flux causes the cores to reach saturation at different times during the next operating half-cycle. During the time interval when one core is saturated and the other is unsaturated, the bridge is unbalanced, giving current flow in the load. The direction of the resultant load current is determined by the polarity of the control signal.

Because of the direction of the rectifiers, conditions during the operating half-cycle are entirely different from those during the reset half-cycle. These two periods of operation must therefore be considered separately.

The Operating Half-Cycle

If output is to be obtained during a given output half-cycle, the bridge must go into this period with a given difference $\Delta\phi$ in the reset flux of the cores. This difference is established by the action of the control windings. If core material having a rectangular hysteresis (B-H) loop and small coercive force is used, practically all the line voltage appears

---

Paper 54-182, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18-22, 1954. Manuscript submitted March 13, 1953; made available for printing December 21, 1953.

C. W. LUFCY and H. H. WOODSON are with the U. S. Naval Ordnance Laboratory, Silver Spring, Md.

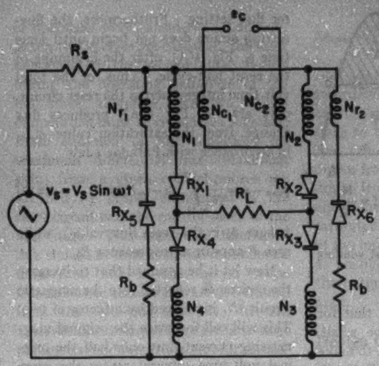


Fig. 1 (left). Half-wave bridge magnetic-amplifier circuit

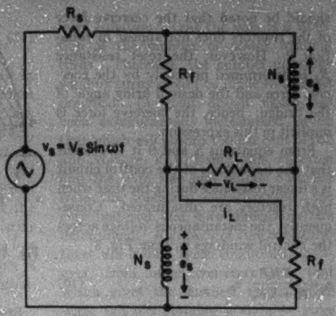


Fig. 2 (right). Equivalent circuit of half-wave bridge with one reactor saturated

across the reactors as long as they are unsaturated. Thus, at the beginning of the output half-cycle, the flux in both cores increases at the same rate until one of the cores saturates. There still remains the initial difference flux $\Delta\phi$ to be changed in the other core before it saturates. During the interval between the time of saturation of the first and second cores, power is delivered into the load. The equivalent circuit of the bridge during this interval is shown in Fig. 2. When the control circuit has been properly designed, by the criterion to be given subsequently, the current in the power winding of the unsaturated reactor is negligible compared to the load current. Thus, by simple voltage division it is seen from this figure that

$$e_s=\frac{R_L+R_f}{R_L+R_s+2R_f}v_s \tag{1}$$

$$v_L=\frac{R_L}{R_L+R_s+2R_f}v_s \tag{2}$$

where

e_s = instantaneous voltage across one unsaturated winding
v_L = instantaneous load voltage
v_s = instantaneous supply voltage
R_L = load resistance
R_f = saturated resistance of one load winding plus forward resistance of its rectifier
R_s = supply source resistance

To simplify this analysis the impedances are assumed to be resistive in nature. For highly rectangular B-H loop core materials, such as Orthonol, the inductive component of the unsaturated as well as the saturated impedance is negligible.[3]

If the time t_1 at which the first core saturates is known, the time t_2 at which the second core saturates is fixed in accordance with Faraday's law, since

$$\Delta\phi=\frac{1}{N_s}\int_{t_1}^{t_2}e_s dt \tag{3}$$

From equations 1 and 3 we have

$$\Delta\phi=\left(\frac{R_f+R_L}{R_s+R_L+2R_f}\right)\frac{1}{N_s}\int_{t_1}^{t_2}v_s dt \tag{4}$$

Since current flows through the load during the interval $t_1<t<t_2$, the average load voltage is

$$V_L=\frac{\omega}{2\pi}\int_{t_1}^{t_2}v_L dt \tag{5}$$

From equations 2 and 5 we have

$$V_L=\frac{\omega}{2\pi}\left(\frac{R_L}{R_L+R_s+2R_f}\right)\int_{t_1}^{t_2}v_s dt \tag{6}$$

Dividing equation 6 by equation 4 and solving for V_L, we obtain

$$V_L=\frac{\omega}{2\pi}\left(\frac{R_L}{R_L+R_f}\right)N_s\Delta\phi \tag{7}$$

This is the gain equation in terms of the equivalent circuit during the operating half-cycle.

From this expression several conclusions may be drawn. First, the gain is proportional to the number of power winding turns N_s. However, it should be pointed out that R_f contains the resistance of the power windings; hence, increasing N_s also increases R_f. R_f also contains rectifier forward resistance. If R_f is small compared to R_L, the gain is relatively unaffected by load resistance. Since R_s does not appear in this expression, the gain is not affected by power source resistance. Although it appears that the gain is a function of frequency, it must be kept in mind that $\Delta\phi$ is a volt-time integral which occurs during the reset half-cycle. In this integral $1/\omega$ is the time scale-factor; consequently, the product $\omega\Delta\phi$ is independent of conditions on the operating half-cycle.

The Reset Half-Cycle

From equation 7 it is seen that for given operating half-cycle conditions maximum gain is obtained when $\Delta\phi$ is a maximum for a given control signal. This signal will come from a control source connected directly to the control windings of the bridge. These control windings present an impedance Z which is determined by the impedances reflected into the magnetic circuit of the reactors. This is shown schematically in Fig. 3.

When the reactors have a rectangular B-H loop core material, the impedance Z, during that portion of the reset half-cycle when both reactors are unsaturated, can be shown to be[4]

$$Z=KN_s^2 \tag{8}$$

with

$$K=\frac{2}{\frac{N_r^2}{R_b}+2\frac{N_s^2}{R_R}} \tag{9}$$

where

N_r = number of reset turns
N_s = number of turns on one power winding
R_b = reset circuit resistance
R_R = reverse resistance of one power-circuit rectifier

It is seen that the factor K is twice the impedance reflected into a 1-turn winding from the reset and power circuits. It

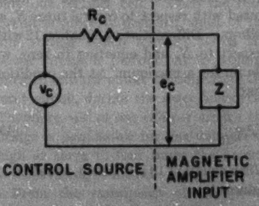


Fig. 3. Equivalent circuit of half-wave bridge input and control source

should be noted that the coercive force of the reactors does not appear in equation 9. However, the reset resistance R_b is determined primarily by the coercive force and the desired firing angle of the bridge; hence, the coercive force is implicit in this expression.

From equations 8 and 9 it is evident that the impedance of the control circuit is a linear resistance during the time when both reactors are unsaturated. Consequently, the instantaneous voltage across the control windings is, from Fig. 3

$$e_c = \frac{KN_c^2}{R_c + KN_c^2} v_c \tag{10}$$

By Faraday's law

$$\Delta\phi = \frac{1}{N_c} \int_{t_3}^{t_4} e_c dt \tag{11}$$

where $t_3 < t < t_4$ is the interval during which the control voltage is producing flux change and $\Delta\phi$ is the flux difference produced in the reactors which it is desired to maximize. For the voltage e_c of equations 10 and 11 to be the same, the interval $t_3 < t < t_4$ must be the interval during which both reactors are unsaturated. In the general case both of the times t_3 and t_4 will vary with control voltage v_c. However, the changes from the values for quiescent conditions ($v_c = 0$) are small; consequently, the error introduced by the assumption that the interval $t_3 < t < t_4$ is constant and equal to the quiescent value is negligible for a sizable range of input signals. The magnitude of this is dependent on the wave form of the control voltage v_c; i.e., if v_c is a pulse occcurring within the interval $t_3 < t < t_4$, this error is zero.

With the assumption that the control interval $t_3 < t < t_4$ is invariant with conditions in the control circuit, substitution of equation 10 into equation 11 yields

$$\Delta\phi = \frac{KN_c}{R_c + KN_c^2} \int_{t_3}^{t_4} v_c dt \tag{12}$$

For a given set of control source conditions the control resistance R_c is fixed; therefore, at any level of control voltage v_c the flux difference $\Delta\phi$ can be maximized with respect to control turns N_c as follows. Differentiating equation 12 with respect to N_c and equating to zero, we obtain for a maximum $\Delta\phi$ the condition that

$$N_c^2 = \frac{R_c}{K} \tag{13}$$

When equation 13 is satisfied, equation 12 reduces to

$$\Delta\phi_{\max} = \frac{1}{N_c} \int_{t_3}^{t_4} \frac{v_c}{2} dt \tag{14}$$

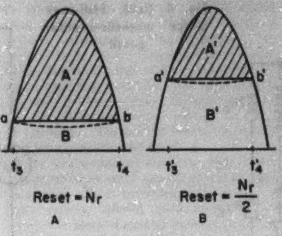


Fig. 4. Effect of reduction of reset winding turns

From equation 14 it is seen that for $\Delta\phi_{\max}$ to be large, N_c should be small. From equation 13 it is seen that to make N_c small, K must be large and R_c small. K is by definition the impedances reflected into the magnetic circuit. These impedances are composed of two components: the resistance of the reset circuit, and the reverse resistance of the power rectifiers. To insure that K be large, it is immediately seen that the coercive force H_c of the core material should be low to maintain a high reset resistance R_b. It is also important that the reset circuit be designed in such a way as to give as high a reflected resistance as is practical. Finally, the reverse resistance of the power rectifiers must be maintained high, commensurate with these and other requirements to be discussed. These resistances which make up K have no theoretical upper limit, although practical limits do exist in that perfect cores and perfect rectifiers are not available. However, from the standpoint of increased gain, a factor in addition to high reflected resistance must be considered for the reset circuit.

The Reset Circuit

When rectangular core material is used, the combined reset and control action does not occur over the entire control half-cycle. This is true since no flux change (control) can occur until a certain level of magnetizing current is reached. Since it falls upon the reset circuit to supply most of this magnetizing current, the values in this circuit will largely determine when, during the cycle, the actual flux change begins to take place. This is illustrated in Fig. 4.

Let it be assumed that to produce a given quiescent firing angle, the reset circuit has N_r turns with a series resistor R_b. Since no flux change occurs until the magnetizing current reaches some value, say i_m, a voltage drop $i_m R_b$ occurs across R_b which is not available across N_r for flux setting. Fruthermore, the flux-setting action does not begin until some time t_3, Fig. 4(A), after the beginning of the reset half-cycle. Thus, of the total volt-time integral across the reset circuit, only area A, Fig. 4(A), produces flux change, from the saturation value $\phi_{\max}$, in accordance with Faraday's law

$$\phi_{\max} - \phi_0 = \frac{1}{N_r} \int_{t_3}^{t_4} (e - i_m R_b) dt \tag{15}$$

(where ϕ_0 is the reset flux value), while area B appears across resistor R_b.

Now let it be assumed that to increase the resistance reflected into the magnetic circuit, N_r is reduced by a factor of two. This will call for twice the original magnetizing current, but only half the original volt-time integral across the reset winding if the same quiescent firing angle is to be maintained. These conditions are met by only slightly decreasing R_b. The aim of increasing the resistance reflected into the magnetic circuit is thus achieved. From Fig. 4(B) it is readily apparent that the control time $t_3' < t < t_4'$ is now less than before. This decrease in time during which control may be effected tends to reduce gain. Thus, reducing reset turns not only increases the reflected resistance, which increases the gain, but also reduces control time, which reduces gain. These two opposite effects must be considered to establish the optimum number of reset turns. Experimental determination of this optimum number for Orthonol cores results in a value of from one-fifth to one-tenth of the total number of power turns on one reactor. This precludes the shunting of power rectifiers to obtain reset through the power windings if optimum operation is desired.

The foregoing discussion of the reset circuit is based on the assumptions that the control interval $t_3 < t < t_4$ is not a function of control voltage and that the leakage through the power rectifiers is negligible compared to the reset circuit current. The first assumption was discussed in the section dealing with the control circuit. The second assumption simplifies the presentation. When the power rectifier leakage current is not negligible, the lower boundaries of areas A and A' in Fig. 4 will be bowed down as shown by the dotted path from a to b in the figure. Even when this occurs the comments relative to reflected resistance from the reset circuit still apply.

Cascading Stages

In the foregoing analysis where only a single bridge was involved, two periods of operation were considered separately. When two bridges are cascaded, the reset half-cycle of the second stage occurs during the operating half-cycle of the first stage; therefore, these two periods of operation must be considered together. From equation 14, the conditions for maximum differential flux (maximum gain) from a given control source are satisfied by proper adjustment of the reset circuit and the control turns of the first stage. This maximum differential flux should in turn establish a maximum differential flux in the second stage. Conditions to achieve this are determined from the following analysis. In this analysis all subscripts 1 are for first-stage parameters and subscripts 2 are for second-stage parameters.

The impedance of the control circuit of the output stage which is the load R_L of the input stage is, by equation 8, therefore

$$R_L = K_2 N_{c_2}^2 \qquad (16)$$

where K_2 is determined by reflected resistances in the magnetic circuit of the output stage and N_{c_2} is the control turns on the output stage.

To achieve maximum control of the second stage, the output interval $t_1<t<t_3$ of the first stage must lie within the control interval $t_3<t<t_4$ of the second stage. (When this condition is satisfied, the control interval $t_3<t<t_4$ of the second stage is determined solely by the reset circuit of this stage.) From Faraday's law

$$\Delta\phi_2 = \frac{1}{N_{c_2}}\int_{t_1}^{t_3} e_{c_2}\,dt \qquad (17)$$

From equations 2 and 16

$$e_{c_2} = v_{L_1} = \frac{K_2 N_{c_2}^2}{K_2 N_{c_2}^2 + R_{s_1} + 2R_{f_1}} v_{s_1} \qquad (18)$$

From equations 17 and 18

$$\Delta\phi_2 = \frac{K_2 N_{c_2}}{K_2 N_{c_2}^2 + R_{s_1} + 2R_{f_1}}\int_{t_1}^{t_2} v_{s_1}\,dt \qquad (19)$$

From equations 4 and 16

$$\Delta\phi_1 = \left(\frac{R_{f_1} + K_2 N_{c_2}^2}{K_2 N_{c_2}^2 + R_{s_1} + 2R_{f_1}}\right)\frac{1}{N_{s_1}}\int_{t_1}^{t_3} v_{s_1}\,dt \qquad (20)$$

From equations 19 and 20

$$\Delta\phi_2 = \left(\frac{K_2 N_{c_2}}{K_2 N_{c_2}^2 + R_{f_1}}\right) N_s \Delta\phi_1 \qquad (21)$$

Maximizing $\Delta\phi_2$ with respect to N_{c_2} (see equations 12 and 13) the condition for maximum gain is obtained

$$(N_{c_2})^2 = \frac{R_{f_1}}{K_2} \qquad (22)$$

From equations 22 and 21

$$\Delta\phi_{2\max} = \frac{1}{N_{c_2}}\,\frac{N_{s_1}\Delta\phi}{2} \qquad (23)$$

If $\Delta\phi_1$ in equation 23 is maximized in accordance with equation 14, maximum gain is obtained from the input of the first stage to the control of the second stage. The same general conclusions may be drawn from equations 22 and 23 as were drawn from equations 13 and 14. From equations 22 and 23 it is seen that the saturated forward resistances of the arms of the input stage should be made as low as possible, and the control windings of the output stage matched to this resistance.

Design Considerations

In undertaking the design of a half-wave bridge magnetic amplifier, the engineer usually has the input or control source characteristics, the output requirements, and the power source given. The load will, by and large, establish the reactor requirements for the output stage. The control source will, as shown in the preceding sections, determine only the number of control turns to be placed upon the input stage cores. Consequently, the design of the amplifier should begin with the output stage and progress to the control source. This procedure is outlined in three steps as follows.

STEP 1

The load requirements for the output stage are generally specified by at least two parameters, from which the load resistance R_L and maximum voltage V_L can be obtained. Available power from the source and allowable amplifier temperature rise will determine the amplifier quiescent power dissipation permissible. From the maximum load current V_L/R_L, the minimum allowable values of wire size and rectifier ratings for the power circuit are determined. Using equation 7 with the given value of R_L, an upper value of R_f is chosen. From this value of R_f, the limit of the forward resistance of the power rectifiers and the resistance of the power windings may be established. In output stages the forward resistance of the rectifiers is usually small compared to the resistance of the power windings. Taking this into consideration, the value of R_f will determine the lengh of wire, of the size chosen, which can be used. The required power source voltage V_s is determined by V_L, R_L, and R_f. If V_s is fixed along with V_L and R_L, there is no choice but to make R_f small enough to satisfy the load requirements. In either case, the maximum length of the power winding is determined by R_f. It may be noted from equation 7 that the gain is a function of the number of power winding turns. Therefore, the number of power winding turns should be made as large as practical. (A good design figure has been experimentally determined to be that which will absorb 150 per cent of full line voltage V_s.) Using this criterion and the wire size and length obtained from the foregoing considerations, the core material and core dimensions are determined. A variety of core configurations will satisfy these conditions; however, from other considerations, that which gives the smallest mean length while still maintaining a suitably large inside-diameter–outside-diameter ratio should be chosen.

From quiescent power dissipation considerations, the firing angle is determined. For this firing angle the reset winding turns are determined to maintain K large. See equations 13 and 14. For Orthonol cores this lies in the range of one-fifth to one-tenth of the total power winding turns on one reactor.

The foregoing procedure establishes the core dimensions, the power circuit, and the reset circuit of the output stage. This completes step 1.

STEP 2

Equation 22 gives the condition to be satisfied by the control turns of the output stage to give maximum gain as indicated by equation 23. At this point K_2 in equation 22 has been fixed by the design values arrived at in step 1. Hence, to achieve maximum gain R_{f_1}, the saturated resistance of an arm of the preceding bridge, must be made an absolute minimum. It is therefore necessary to determine how low R_{f_1} can be made before setting N_{c_2}, the number of control turns on the output stage.

The lower limit on the value of R_{f_1} is determined by rectifier forward resistance and power winding resistance. From the standpoint of low forward resistance of the rectifiers in this application, the best that can be done is to use only one rectifier cell. When the reset circuit is properly designed, most of the line voltage which appears across the power circuit is bucked out by voltage induced into the power windings from the reset circuit during the reset half-cycle. Therefore, use of one cell per arm of the bridge is permissible for line voltages as high as three to four times the inverse voltage rating of the cell.

From the standpoint of the power windings, the lower limit on R_{f_1} is set by the increase in quiescent power and core size which results if larger wire is used.

Increase in quiescent power can, however, be controlled by the insertion of a line resistor in series with the power source. Equation 7 shows that the inclusion of such a resistance does not affect the gain of the stage.

Once R_{f_1} is determined, N_{c_1} is designed in conformance with equation 22. In arriving at the minimum values of R_{f_1}, the core size and power circuit of the preceding stage will have been determined on the basis of considerations outlined in step 1. Given the core size and power circuit for this stage, the design of the reset circuit is carried out in the same manner as for the output stage. This completes step 2.

Up to this point the output stage has been completely designed as have the core size, power circuit, and reset circuit of the preceding stage. If this preceding stage is not the input, step 2 is repeated until the input stage is reached.

STEP 3

The input stage control windings are designed in accordance with equation 13 which matches the control windings with the control source, given in the initial requirements. This completes step 3 and the amplifier design.

Additional Considerations

RECTIFIERS

It is obvious that for magnetic-amplifier applications a perfect rectifier, i.e., zero forward resistance and infinite back resistance, is desirable because it would be applicable in all cases. In practice, however, lower forward resistance is usually obtainable only with attendant lowering in back resistance and, conversely, increased back resistance is obtainable only with increased forward resistance. Therefore, it is seen from the foregoing discussion that the principal requirements of a specific application must be investigated before rectifiers having the most suitable properties can be chosen.

CORE MATERIAL

From the foregoing analysis it is clear why the core material for use in this circuit should have a high saturation flux density, a high degree of rectangularity of the B-H loop, and as low a coercive force as possible. First, for a given core size, a high saturation flux density allows large values of supply voltage. This gives increased power handling capacity.

It is of utmost importance that the unsaturated impedance of the reactor power winding be very high while its saturated impedance be as close to the winding resistance as possible. The more rectangular the B-H loop of the core material, the more closely this condition is approached.

The lower the coercive force of the core, the higher the value of K which is defined by equation 9. Higher values of K result in lower values of N_c (see equation 13) with an attendant increase in gain as shown by equation 14. Of the two factors that determine K, namely the reflected resistance of the reset circuit and the back resistance of the rectifiers reflected through the power windings, the first varies inversely with the coercive force of the core without limit. The second, however, will never be greater than the back resistance of the rectifiers regardless of the coercive force of the core. The quality of dry rectifiers now available is such that their low back resistance limits K to the extent that reduction in coercive force of the core material which usually is obtained at the expense of rectangularity of the core B-H loop is a questionable approach to achieving increased gain. Thus, improvement in amplifier gain can be made only by simultaneous improvement in rectifiers and core materials.

TIME DELAY

One of the main advantages of the half-wave bridge magnetic amplifier is its fast speed of response which is independent of gain. When properly adjusted, the circuit has a half-cycle per stage delay which is the fastest response time possible in a magnetic amplifier. Certain precautions must be exercised, however, to insure that the full capabilities of the bridge are retained.

If the power windings are designed to absorb more than full line voltage it is possible to have a delay greater than one half-cycle. This condition is illustrated in Fig. 5. Under normal quiescent conditions the cores in the bridge are cycling about a minor loop A, Fig. 5(A). When only sufficient control is applied to cut off one of the cores, the flux change during the reset half-cycle must be exactly equal to the flux change during the operating half-cycle. The core is then operating about loop B, Fig. 5(A). When additional control is applied, the flux change during the reset half-cycle becomes greater than that during the operating half-cycle, hence the flux in the core cycles toward negative saturation as illustrated in Fig. 5(B). When this control is removed the flux will cycle back up the loop until it reaches the normal quiescent loop A, Fig. 5(A). This cycling back up the loop entails a delay of at least one and possibly more cycles of the supply frequency. Although this effect gives no additional delay in output rise time, it does cause a definite increase in the fall-off time. There are three ways to prevent this. The first is to design the reactor so it will not absorb more than line voltage. This is done at the sacrifice of gain. Another way is to keep the power source impedance low. Inspection of Fig. 2 will show that if R_s is low, considerably more than half the line voltage appears across only half the power windings. This will tend to insure saturation of the core. Finally, if a control source is used which is incapable of driving the reactor to cut off, this delay will be avoided.

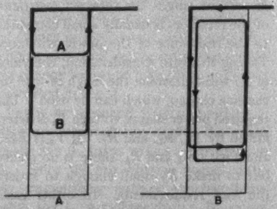


Fig. 5. Effect of control on reactor flux

If the load is highly inductive, the load current will have a finite decay time. If this decay time runs into the following reset half-cycle, this current behaves as an additional control on the bridge. This results in an increase in the response time of the amplifier.

When two stages, each having a half-cycle response, are coupled together, the combination will sometimes show a response time longer than should be expected. This is due to the induced voltage which appears in the control winding of the second stage when it is delivering output. This voltage causes control action on the first stage through leakage in the power rectifiers of the first stage. Good rectifiers in the first stage will reduce this action to a minimum.

DRIFT

Since the half-wave bridge magnetic amplifier is a balanced circuit, drift is merely a function of parameter balance on each side of the bridge. The only way drift can be eliminated is to maintain this parameter balance under all conditions of temperature, voltage, frequency, etc., to be encountered. To do this requires careful matching of components. The degree of care which must be put into the matching of components is determined by the minimum allowable drift.

Both rectifiers and cores are fairly easy to match for variations in line voltage and frequency at any given temperature. The temperature drift problem, however, is not so readily overcome. Dry rectifiers which are presently available show an extremely wide variation in characteristics with temperature, and these variations are frequently not reproducible unless the units have been temperature-cycled and aged by some voltage-temperature schedule designed to stabilize them. One common approach to the rectifier drift problems is to check single cells and group those cells with similar characteristics. From these groups of similar cells, matched rectifier sets can be assembled. In spite of the difficulties, rectifiers have been matched on a practical basis to the point where temperature drift due to their unbalance is less than 10^{-6} watts of input power, for a temperature range of -55 to $+60$ degrees centigrade.

Unmatched cores also contribute substantially to the temperature drift. Here the problem is more serious. The drift from the cores is probably due primarily to expansion and contraction of the core which affects the strain-sensitive properties of B-H loop rectangularity and coercive force of the core material. From this standpoint, the toroidal tape configuration could be inferior to a laminated core. It may well be that in certain extremely critical applications a laminated core structure in the input stage may be superior to the tape toroid for minimizing temperature drift. Matching at room temperature is not sufficient for critical applications. Fortunately, if cores are carefully insulated and heat-treated, changes brought about in them by temperature variations are similar. Thus drift which is caused by differential changes will be small in well-matched units.

Conclusions

From the theoretical considerations of the operation of the half-wave bridge, a logical qualitative design procedure has been set forth which greatly simplifies the design of an amplifier having optimum characteristics. Since the foregoing design procedure is qualitative, good engineering judgment is required to obtain actual design figures. Quantitative design data can be obtained by a more thorough mathematical derivation based on the considerations given here. Such a derivation has been carried out at the U. S. Naval Ordnance Laboratory.[4]

References

1. Compensation of a Magnetic Amplifier Servo System, H. H. Woodson, A. E. Schmid, C. V. Thrower. *Proceedings*, National Electronics Conference, Chicago, Ill., 1952.

2. An Improved Magnetic Servo Amplifier, C. W. Lufcy, A. E. Schmid, P. W. Barnhart. *AIEE Transactions*, vol. 71, pt. I, 1952, pp. 281–89.

3. A Non-Linear Resistive Element, C. M. Davis, Jr. *NavOrd Report 2700*, U. S. Naval Ordnance Laboratory, Silver Spring, Md., Dec. 30, 1952.

4. Preliminary Analysis of the Half-Wave Bridge Magnetic Amplifier, H. H. Woodson. *NavOrd Report 3596*, U. S. Naval Ordnance Laboratory, Silver Spring, Md., Dec. 1953.

◆

Discussion

R. O. Decker (Westinghouse Electric Corporation, East Pittsburgh, Pa.): The authors, in making an analysis of the half-wave bridge circuit, have evidently assumed the cores to have rectangular hysteresis loops with fixed vertical sides. The result is that the equations developed do not show clearly the dependence of the power gain of the amplifier upon the magnetic characteristics of the cores.

If the quantity $V_c=(\int_0^T v_c dt)/T=f\int_{t_3}^{t_4} v_c dt$ is considered to be the average electromotive force of the control source, and if equations 12 and 7 are combined, the power gain of the amplifier will be

$$A_p=\left(\frac{V_L}{V_c}\right)^2\frac{R_c}{R_L}=\left(\frac{R_L}{(R_L+R_f)^2}\right)N_s^2\times\left(\frac{\dfrac{N_c^2}{R_c}}{\left(\dfrac{1}{K}+\dfrac{N_c^2}{R_c}\right)^2}\right) \qquad (24)$$

If the condition for maximum voltage gain is now imposed, i.e., $N_c^2/R_c=1/K$, the maximum power gain of the amplifier is

$$A_p=\left(\frac{R_L}{(R_L+R_f)^2}\right)\left(\frac{N_s^2}{4}\right)K \qquad (25)$$

If rectifier leakage is neglected, the expression for power gain, using the authors' value for K, is

$$A_p=\left(\frac{R_L}{(R_L+R_f)^2}\right)\left(\frac{N_s^2}{2}\right)\left(\frac{R_B}{N_r^2}\right) \qquad (26)$$

The authors state that the coercive force of the core is implicit in the reset resistance R_B, but suppose that the reset voltage is supplied from an external high-voltage source. Under these conditions R_B could be made a very large resistance and it might be concluded, from the preceding equation, that the power gain would vary directly as R_B and that no upper limit on the power gain would exist. Experimentally, this is not found to be the case.

A more logical expression for power gain can be developed if some acknowledgment is made of the magnetic characteristics of the cores. I have made an analysis which results in the following expression for the maximum power gain of the single-stage amplifier

$$A_p=\left(\frac{R_L}{(R_L+R_f)^2}\right)\frac{N_s^2}{2}\left(\frac{1}{\dfrac{1}{K_c}+\dfrac{N_r^2}{R_B}}\right) \qquad (27)$$

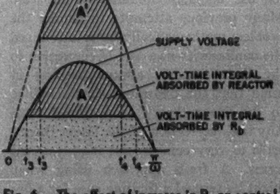
This expression for maximum power gain, which has been checked experimentally, contains a term that is descriptive of the magnetic properties of the core. The factor K_c relates the flux levels in the amplifier to the dynamic magnetic characteristics of the cores, and it will be noted that the upper limit on the power gain of the amplifier is determined by this relationship. In applications where R_B/N_r^2 is of the same order of magnitude as K_c, each term is important in determining the maximum power gain of the amplifier.

C. W. Lufcy and **H. H. Woodson:** We make clear in the paper that a condition for the validity of the analysis is that a rectangular core material be used. (To be rectangular the B-H loop must have "fixed vertical sides.") Specifically, just after equation 2, immediately prior to equation 8, and at the beginning of the section on the reset circuit is the point stressed that rectangular core material is assumed.

Mr. Decker states that, if rectifier leakage is neglected, the power gain varies directly with the reset resistance, and therefore one might conclude that if the reset is supplied from a high-voltage source, the reset resistance can be increased with a consequent increase in power gain. The fallacy in this argument is in the determination of the limits of integration on the term that he calls the average input voltage

$$V_c=f\int_{t_3}^{t_4} v_c dt$$

This can be shown in the same manner as in the section on the reset circuit in the paper. First, assume that for a specific quiescent firing angle the conditions are as shown in A in Fig. 6. Double the voltage supplied to the reset circuit will be assumed. If the reset circuit turns are unchanged, to still have the same volt-time integral applied to the reset winding, the reset resistance



A'
A
SUPPLY VOLTAGE
VOLT-TIME INTEGRAL ABSORBED BY REACTOR
VOLT-TIME INTEGRAL ABSORBED BY R_b
0
t_3
t_3'
t_4'
t_4
$\frac{\pi}{\omega}$

Fig. 6. The effect of increase in R_b on control time, with external high-voltage source

must be increased, giving conditions as shown in A' in Fig. 6. At the same time, however, the control interval, which was $t_3<t<t_4$, is now reduced to $t_3'<t<t_4'$. If this process is carried further the effect of the increase in R_b must be balanced against the decrease in control time to determine the variation in gain. Since the control source is supplying power even when the reactor is not controlling $(0<t<t_3;\ t_4<t<\pi/\omega)$, it seems unfair to include only that control power supplied during the time while flux is actually being changed in the reactors in defining V_e.

From this it is evident that the theory given in the paper predicts that the power gain will vary in a manner indicated by Mr. Decker's experimental results.

Mr. Decker has simplified the expression for power gain derived from the paper under consideration by neglecting rectifier leakage. It is pointed out that in the paper itself no such assumptions are made. The authors know of no commercially available dry rectifiers with negligible back leakage under all conditions encountered in the magnetic-amplifier circuit under consideration.

If the value for K given in equation 9 is substituted into the expression for power gain which Mr. Decker gives, we obtain

$$A_p=\left(\frac{R_L}{(R_L+R_f)^2}\right)\frac{N_s^2}{2}\left(\frac{1}{2\dfrac{N_s^2}{R_B}+\dfrac{N_r^2}{R_B}}\right) \qquad (28)$$

This is very similar to equation 27 derived by Mr. Decker.

To properly compare and evaluate Mr. Decker's expression for power gain it would be necessary for us to know how the factor K_e is defined and/or determined. Indeed, it is possible that his value $1/K_e$ is nothing more than our value $2N_s^2/R_B$ in disguise (or vice versa).

Eddy-Current Phenomena in Ferromagnetic Materials

H. M. McCONNELL
ASSOCIATE MEMBER AIEE

INTEREST in eddy currents in solid iron masses has kept pace with the development of electromagnetic devices generally.[1] The first problems arose in the design of eddy-current brakes for flywheels. Then the use of iron wires for telephone lines and iron rails for the supply of power to a-c locomotives led to new problems. It was recognized that the saturation of the iron was an important factor, and many authors have presented their theories to take saturation into account. Following the appearance of Rosenberg's work,[2] others[3–6] have made an academic problem of substantiating it or elaborating upon it. Recently the damping effect of eddy currents in the solid yokes and pole pieces of d-c machines has become important. Still more recently, it has been recognized that the saturation effect is important in computing the core losses in thin steel or alloy sheets.

A new method for computing the effect of saturation has been suggested.[7–9] The method seems to be a significant departure from other efforts to account for the magnetic nonlinearity in iron. Many previous attempts to handle the problem mathematically have introduced the nonlinearity as a sort of correction to the linear theory, always keeping in mind that the treatment should reduce to the linear theory as a special case. Any such correction of the linear theory leads to very cumbersome mathematical forms. A fresh approach using what might be termed a limiting case for a beginning may, if carefully applied, lead to a mathematical formulation even simpler than in the linear case. It is the purpose of this paper to show a logical transition from the linear theory of eddy currents in solid media to the limiting non-linear theory, and to show the applicability of the method in typical problems.

An almost exact parallel can be found in the various approaches to the nonlinear problem of the magnetic amplifier. On the one hand, there is the formulation by fitting some useful function, such as the hyperbolic sine or a finite power series, to the magnetization curve. On the other hand, a radically different approach exists in which all resemblance to linear behavior of the magnetic material is discarded. These various analyses of the magnetic amplifier are outlined and compared in reference 10.

Comparison of Linear and Limiting Nonlinear Theories

The linear theory will be reviewed by the use of a simple situation. An infinite half-space with its surface in the x, y plane is excited such that the total magnetic flux carried in the x direction per unit of width in the y direction is $Re\phi_{\max}\epsilon^{j\omega t}$ There is no y or z component of flux. The z direction extends into the material normal to the x, y plane. The permeability of the material is μ, its conductivity is σ, and the radian frequency is ω. Taking **B**, the induction, to be the field variable, the field equations are $B_y=0$, $B_z=0$, and (using meter-kilogram-second units)

$$\frac{\partial^2 B_x}{\partial z^2}=\mu\sigma\frac{\partial B_x}{\partial t} \qquad (1)$$

The solution is

$$B_x=ReB_0\epsilon^{(j\omega t-z\sqrt{j\omega\mu\sigma})} \qquad (2)$$

where B_0, a complex number, is the induction at the surface. The total flux per unit width is

$$\phi(t)=\int_0^\infty B_x dz=\frac{B_0}{\sqrt{j\omega\mu\sigma}}\epsilon^{j\omega t} \qquad (3)$$

Substituting into equation 2

$$B_x=Re\sqrt{j\omega\mu\sigma}\phi_{\max}\epsilon^{(j\omega t-z\sqrt{j\omega\mu\sigma})} \qquad (4)$$

Expressed in a different form equation 4 becomes

$$B_x=Re\frac{\phi_{\max}}{\dfrac{1}{\sqrt{\omega\mu\sigma}}}\epsilon^{-z\sqrt{\omega\mu\sigma/2}}\epsilon^{j(\omega t+\pi/4-z\sqrt{\omega\mu\sigma/2})} \qquad (5)$$

where $\phi_{\max}$ is a real number.

The behavior of B_x as a function of time and of depth is illustrated in Fig. 1. The upper curve is the total flux $\phi(t)$ while the lower curves give the flux density as a function of time at the various levels. At any instant, the flux density varies as a damped sinusoid with depth; thus there are alternate positively and negatively magnetized bands of material. The existence of these alternate bands can be seen also by inspecting the time function at various levels.

Now the same situation is analyzed once more, but with the linear magnetic material being replaced by a material having the nonlinear magnetic characteristic shown in Fig. 2. This material is magne-

---

Paper 54-146, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 20, 1953; made available for printing December 7, 1953.

H. M. McCONNELL is with the Carnegie Institute of Technology, Pittsburgh, Pa.

This work was supported in part by the Office of Naval Research, under Contracts *N7 ONR 30306* and *30308*, projects *975-272* and *275*.

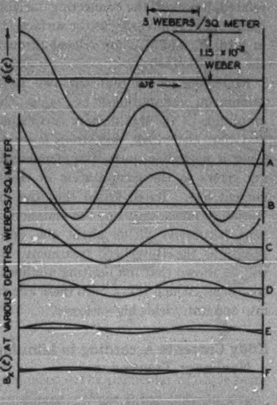


Fig. 1. Flux per unit width and flux density in the classical theory, with constant permeability

$\omega = 377$
$\mu\sigma = 6.2 \times 10^4$

Curve	Width, z, Millimeters
A	0
B	0.176
C	0.352
D	0.528
E	0.704
F	0.880

This table also applies to Fig. 4.

tized to saturation if the field intensity is different from zero, and it is possible to change the flux density only at H=0. The statement implies that if the flux density is changing, H must be zero. However, the converse statement is not necessarily true; if H is zero, the flux density is not necessarily changing. Thus it is possible to have regions within this material where H is zero but where the flux density can have any constant value less than or equal to the saturation induction. The particular constant value would depend upon the state in which the material was left during some previous process.

The mechanics of supporting a function $\phi(t) = Re\phi_{max}\epsilon^{j\omega t}$ will now be investigated. First it will be assumed that H at any point is a periodic function with fundamental radian frequency ω and no average value. If this function has zeroes only at discrete values of (ωt) which are π radians apart, the corresponding induction B will be a square wave of period $2\pi/\omega$ and amplitude B_s. On the other hand, a square wave of B at any point can be supported by a discontinuous function $H(\omega t)$, due to the possibility (discussed before) that B may remain constant when H is zero. Thus it seems reasonable to postulate that the induction at any point in the material is either constant at some value depending on previous treatment or a square wave with amplitude B_s and period $2\pi/\omega$.

Inspection of Fig. 1 gives some insight as to how the square waves giving B as a function of time might combine to give a total flux $\phi(t)$ which is sinusoidal. In Fig. 1, the induction at any level z is a sine wave versus time, but whose phase is shifted with respect to the level above. The next question is, can a succession of square waves which are phase-shifted with respect to one another add to a sine wave? Fig. 3 shows a finite group of these square waves and their instantaneous sum. (The resemblance to the calculation of the magnetomotive force of a distributed winding in a rotating machine is apparent.) If there are enough square waves, phase-shifted in a particular manner which can be computed, a smooth sine wave $\phi(t)$ results.

The computation of this phase shift as a function of depth is better approached by a change of variable. At some instant, say $(\omega t)_1$, the induction above the level z_1(Fig. 3) is $-B_s$ while below that level it is $+B_s$. At a later instant $(\omega t)_2$ the induction above the level z_2 has become $-B_s$ while below z_2 it remains $+B_s$. Thus a surface of separation between the two saturated states has moved from z_1 to z_2 during the interval (t_2-t_1). The phase shift between the square waves of induction at these two levels will be $\omega(t_2-t_1)$ radians. The distance (z_2-z_1) that the surface of separation has moved, and the phase shift $\omega(t_2-t_1)$ radians, are both related to the change in flux per unit width $\Delta\phi$ which has occurred during that time.

The surface of separation will stop when there is no further need to "substract flux" by causing this change to take place. The maximum distance that this surface travels, beginning at $z=0$, will be the depth of penetration δ. When the surface of separation has reached the depth $z=\delta$, the flux per unit width will be $-\phi_{max}$. The depth of penetration is

$$\delta = \frac{\phi_{max}}{B_s} \tag{6}$$

This movement from $z=0$ to $z=\delta$ has occurred in a half-cycle, since the flux per unit width has been changed from ϕ_{max} to $-\phi_{max}$. Thus, the phase shift between square waves of B versus (ωt) at $z=0$ and $z=\delta$ is π radians.

The depth of penetration δ is seen to be a variable depending on ϕ_{max}, while in the linear material the depth of penetration is a property of the material at a

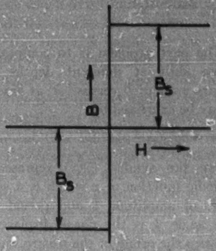


Fig. 2. Assumed magnetic characteristic

given frequency (equation 4). This behavior is expected, since in the linear material it is possible to increase $B_{z\,max}$ if more flux is required; in the nonlinear material the induction is limited to B_s, and more flux can be obtained only if δ increases.

The position of the surface of separation will now be determined as a function of the flux ϕ. Let the location of the surface be z'. The flux per unit width contributed when the surface moves from z' to $(z'+\Delta z')$, changing the state from $+B_s$ to $-B_s$, will be

$$\Delta\phi = -2B_s\Delta z' \tag{7}$$

Passing to the limit $\Delta\phi \to 0$ yields the differential equation for z'

$$\frac{dz'}{d\phi} = -\frac{1}{2B_s} \tag{8}$$

The integration will be performed beginning at the time when $\phi = +\phi_{max}$ and $z'=0$. (Calculations begun when $\phi = -\phi_{max}$ would proceed in the same way with a change in sign, such that z' is always a positive increasing function). Thus

$$z'(\phi) = \frac{1}{2B_s}\int_{\phi_{max}}^{\phi} -d\phi \tag{9}$$

or

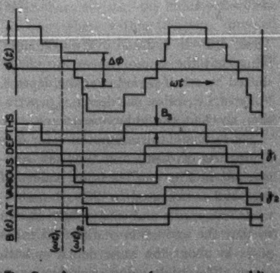


Fig. 3. Arrangement of square waves adding to a smooth periodical function

$$z'(t)=\frac{\phi_{\max}-\phi(t)}{2B_s} \tag{10}$$

The next problem is to compute the phase shift at different depths. The square wave of B versus (ωt) at any level z may be expressed by

$$B(\omega t)=B_s\square(\omega t-\psi) \tag{11}$$

where the symbol $\square(x)$ represents a periodic function of x which is a square wave of amplitude 1 and period (2π). The angle ψ is the phase displacement between the origin of the square wave at depth z and the origin of the square wave at $z=0$, so ψ will be a function of z. The origin of the square wave is taken at the center of the positive flat portion.

Equation 10 was derived assuming that at time zero the flux is a positive maximum and is being decreased. Thus, at time zero all the material was previously magnetized to $+B_s$ induction, and will be changed during the subsequent half-cycle to $-B_s$ induction. At $z=0$ the origin of the square wave would be located at $\psi=-\pi/2$, since at time zero that value of $(\omega t-\psi)$ must be $\pi/2$. The further dependence of ψ upon z can be deduced from equation 10, since when the surface of separation is located at $z'=z$ we must have $(\omega t-\psi)=\pi/2$ at that level. Thus setting

$$z'\left(\frac{\pi/2+\psi}{\omega}\right)=z$$

yields $$z=\frac{\phi_{\max}-\phi\left(\frac{\pi/2+\psi}{\omega}\right)}{2B_s} \tag{12}$$

If now $\phi=Re\phi_{\max}e^{j\omega t}$ is substituted in equation 12, the relationship between ψ and z for that case is

$$z=\frac{\phi_{\max}}{2B_s}\left[1-\cos\left(\frac{\pi}{2}+\psi\right)\right]$$

$$=\frac{\delta}{2}[1+\sin\psi] \tag{13}$$

or

$$\sin\psi=2\frac{z}{\delta}-1 \tag{14}$$

Thus, $\psi(0)=-\pi/2$ and $\psi(\delta)=\pi/2$, signifying that the square waves are phase-shifted by a half-period over the depth δ. For $z>\delta$, i.e., below the depth of penetration, no change in induction occurs.

The results of this computation for B as a function of depth and time are summarized in Fig. 4. The same values of $\phi_{\max}$ and ω are used in Figs. 1 and 4, while values of μ and σ have been chosen so that the extinction of the field takes place in about the same depth in both cases. The peak value of induction is seen to be much higher in Fig. 1 than in Fig. 4, while the phase shift is larger in Fig. 4. The different phase shift might have been anticipated; in the linear case, Fig. 1, the contribution to $\phi(t)$ by the lower levels, where the phase is advanced, has been reduced due to the damping effect. In the nonlinear case this gradual damping does not exist. This comparison also validates the idea of using square waves of flux; the saturation induction is reached very early in the cycle in Fig. 1.

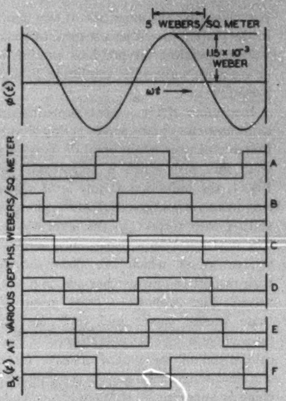


Fig. 4. Sinusoidal flux per unit width in saturated solid ion, and square waves of flux density at various depths. See Table of Fig. 1

$\omega=377$
$B_s=1.3$
$\sigma=6.67\times10^8$

An Intermediate Theory

A theory to account for saturation, which in a sense is intermediate between the linear theory and the limiting nonlinear theory, has been proposed by Barth.[3] It was assumed that the flux density varies sinusoidally with time at any level, as in the linear theory, but that the amplitude of the sine wave is B_s at every level (except possibly in that part of the material which experiences no change in induction). A detailed review of this intermediate theory is presented in reference 11.

The Theory of Rosenberg

Rosenberg's treatment[2] assumes that the flux density is a sinusoidal function, without phase shift from layer to layer, and with a constant amplitude $B_{\max}$ between $z=0$ and $z=a$. The amplitude $B_{\max}$ and the depth of penetration a are related, however, the connecting variable being the field intensity at the surface H_0; in turn, H_0 and $B_{\max}$ are related through the saturation curve.

Rosenberg noted that the two most prominent assumptions, i.e., that the amplitude of the flux density is constant and that the phase shift from layer to layer is negligible, introduce compensating errors in the computation of eddy-current loss. Barth found that the inclusion of phase shift but the omission of time harmonies yields about the same results as the Rosenberg treatment. It will be shown that the limiting nonlinear treatment, which takes both these effects into account, yields higher losses.

Eddy Currents According to Limiting Nonlinear Treatment

A basic assumption in this treatment is that a change in induction can take place only when $H=0$. Therefore, the sum of all the field intensities at the surface of separation between the states $-B_s$ and $+B_s$ must be zero. This statement allows the derivation of a relationship among the flux, the externally applied field, and the eddy currents.

The field equations which relate the density of eddy currents to the rate of change of magnetic induction are

$$\mathbf{J}=\sigma\mathbf{E} \tag{15}$$

$$\text{curl}\,\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t} \tag{15A}$$

Equation 15(A) says that the electric field intensity $\mathbf{E}$ is everywhere irrotational, because the induction can be only $\pm\mathbf{i}B_s$ where $\mathbf{i}$ is the unit vector in the x-direction. The only possible irrotational vector $\mathbf{E}$ under the conditions set for the problem is a vector varying in time but not varying with position at a given instant. To find out what $\mathbf{E}$ is, it is necessary to convert the equation 15(A) to the integral form, yielding

$$\mathbf{J}=\sigma\mathbf{E}$$

$$\oint_C\mathbf{E}\cdot d\mathbf{r}=-\frac{\partial}{\partial t}\int\int_S\mathbf{B}\cdot d\mathbf{S} \tag{16}$$

where C bounds $\mathbf{S}$ and traverses $\mathbf{S}$ in the usual counterclockwise sense.

The line integral in equations 16 will be zero if C lies entirely on one side of the surface of separation between $-B_s$ and $+B_s$. However, if C cuts this surface, then the integral over $\mathbf{S}$ will have a time derivative depending upon the velocity of the surface of separation and the length of the segment common to both $\mathbf{S}$ and the moving surface. Thus a discontinuity in $\mathbf{E}$ is expected at the moving surface.

The integrals in equations 16 are easily

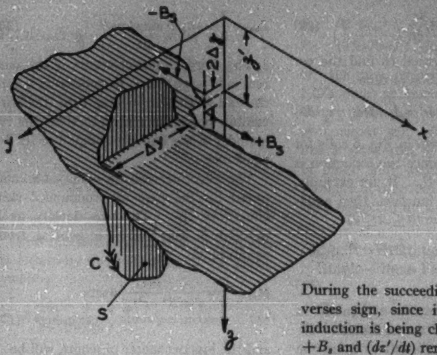


Fig. 5. Definition of symbols used in plane configuration

evaluated in the particular example being considered. Since **B** has only an x-component, the surface **S** is taken in the y,z plane. The surface of separation between states $-B_s$ and $+B_s$ is an x,y plane moving in the z-direction. The segment common to both these surfaces will be Δy meters long. The electric field **E** will have a y-component only, which will (at any instant) take constant values above and below the surface of separation, with a discontinuity at the surface. The situation is illustrated in Fig. 5.

Let the induction (in the x-direction) be $-B_s$ above the surface and $+B_s$ below, with the surface moving at velocity (dz'/dt) meters per second. At time t, let **S** have S_1 area in the region of $-B_s$ induction and S_2 area in the region of $+B_s$ induction. At time $(t+\Delta t)$ the surface will have moved a distance $(dz'/dt)(\Delta t)$ meters, so that now there will be $[S_1+(dz'/dt)(\Delta t)(\Delta y)]$ area in the region of $-B_s$ induction and $[S_2-(dz'/dt)(\Delta t)(\Delta y)]$ area in the region of $+B_s$ induction. The integration in equations 16 becomes

$$(E_y|_{z'-\Delta z})(\Delta y)-(E_y|_{z'+\Delta z})(\Delta y) = \frac{2B_s(dz'/dt)(\Delta t)(\Delta y)}{\Delta t} \qquad (17)$$

or

$$(E_y|_{z'-\Delta z})-(E_y|_{z'+\Delta z})=2B_s\frac{dz'}{dt} \qquad (18)$$

Equation 18 gives the discontinuity in the electric field existing at the moving surface separating $-B_s$ and $+B_s$.

There can be no electric field below the moving surface, for if it were to exist energy would be required to sustain the resulting eddy currents and there is no such energy available. Therefore

$$E_y=2B_s\frac{dz'}{dt}\ (0\leq z<z');(B_x|_{z'-\Delta z}=-B_s) \qquad (19)$$

$$E_y=0 \qquad (z'\leq z)$$

During the succeeding half-cycle E_y reverses sign, since in that removal the induction is being changed from $-B_s$ to $+B_s$ and (dz'/dt) remains positive.

The eddy currents flow with uniform density (at any instant) in the space $(0\leq z<z')$ and there are no eddy currents below the level z'. The total current contained in the eddy-current path is, by equations 15 and 19

$$I_e=2B_s\sigma z'\frac{dz'}{dt}\ \text{amperes per meter} \qquad (20)$$

during the half-cycle in which the material is being changed from $+B_s$ to $-B_s$. During the following half-cycle, I_e changes sign. Since the field intensity must be zero at the surface between $-B_s$ and $+B_s$, the applied field intensity must be the opposite of equation 20; in other words, the applied field must be $-I_e$ amperes per meter.

An expression for the applied field J_s can be obtained from equations 20 and 10; i.e., for sinusoidal flux variation

$$J_s=-2B_s\sigma\left(\frac{\phi_{\max}-\phi(t)}{2B_s}\right)\left(-\frac{1}{2B_s}\frac{d\phi}{dt}\right) = \frac{\sigma}{2B_s}(\phi_{\max}-\phi)\frac{d\phi}{dt} \qquad (21)$$

In the present case, $\phi=\phi_{\max}\cos\omega t$ so that during the half-cycle when ϕ proceeds from $\phi_{\max}$ to $-\phi_{\max}$

$$J_s=-\frac{\omega\sigma}{2B_s}\phi_{\max}^2(1-\cos\omega t)\sin\omega t\ \text{amperes per meter} \qquad (22)$$

Equation 22 is similar to equation 9 of reference 6. The wave form of J_s is repeated during the following half-cycle but with reversed sign (it is convenient to establish a new time origin each half-cycle).

Development of the Nonlinear Treatment

Rosenberg's treatment of eddy-current loss in solid iron was the first attempt of practical value to account for the effect of saturation. It was based on an hypothesis which fitted quite well with the over-all test results then known. However, the essential nature of the field problem was not considered, since Rosenberg neglected both the phase shift in the field from layer to layer and the fact that the induction at any depth must have prominent time harmonics. As a result both Rosenberg's and Barth's treatments yield a sinusoidal flux for a sinusoidal current, which is known to be in error, especially when core materials are used whose characteristics approach Fig. 2. It has been shown here that the limiting nonlinear treatment takes into account both the phase shift and the time harmonics. The test results reported by Hale and Richardson[8] confirm its validity.

Applications of Limiting Nonlinear Treatment

Applications of this new theory have been restricted so far to the computation of eddy-current loss in thin sheets[6] and the explanation of certain effects in pulse transformer cores.[9] To illustrate the versatility of the theory, it will be used in two other problems of continuing interest. These are the computation of losses in solid iron conductors, and the design of inductive heating apparatus for use on a charge of solid iron.

Losses in Iron Wire

The first problem is to deduce the field configuration within the wire under the present assumptions. The current will flow in the axial direction, and the wire will become magnetized in a circumferential sense. Suppose for the moment that a steady direct current flows in the wire. The current density will be uniform, and application of Ampere's law shows that H has a value everywhere within the wire, being zero at the center only. According to the assumed magnetic characteristic, the induction will be B_s everywhere within the wire, directed circumferentially. Now if the direct current is reduced to zero the wire will remain in this magnetic state.

Next suppose that a direct current of opposite sign is to be established. The new steady magnetic state will be at saturation induction, but in the opposite direction. The change in induction can occur only if $H=0$. Ampere's law leads to the conclusion that there must be a circular region within the wire which carries no current, such that $H=0$ at the boundary of this region; the induction is the old value of, say, $-B_s$ in that region. The new current to be established is forced to flow outside this region. As the new current builds up, the region of $-B_s$ is forced to shrink toward the center, and after a time complete reversal of the induction within the wire may be accomplished.

Fig. 6 shows the wire, with the boundary between states $+B_s$ and $-B_s$ located at $r=r'$. This moving boundary causes a change in the flux linkage with that part of the wire within the boundary, i.e., $(r<r')$, while the part of the wire in the region $(r'<r<R)$ sees no change in flux linkage (except the change in the field outside the wire, which is common to both regions and need not be considered). Since the axial electric field is the same in both regions, the resistance drop in the region carrying current can be equated to the drop due to the induced electromotive force in the region carrying no current.

Right-handed cylindrical co-ordinates are used, Fig. 6. At time t, the surface of separation is located at $r=r'$. The electric field in volts per meter E_z, impressed along the element whose area is dS, is to be computed. The flux linkage with that section at time t is

$$\lambda(t)=-B_s(r'-r_s)+B_s(R-r')\text{ webers per meter} \tag{23}$$

At time $(t+\Delta t)$, the surface has moved to $(r'+\Delta r')$ and the flux linkage has changed to

$$\lambda(t+\Delta t)=-B_s[(r'+\Delta r')-r_s]+B_s[R-(r'+\Delta r')]\text{ webers per meter} \tag{24}$$

On passing to the limit, it is found that

$$\frac{d\lambda}{dt}=-2B_s\frac{dr'}{dt} \tag{25}$$

which is independent of r_s. Therefore every element dS within the region $(0<r<r')$ sees the same electric field impressed, which is

$$E_z=-2B_s\frac{dr'}{dt}\text{ where }(0\leqslant r<r') \tag{26}$$

Outside this region the electric field E_z is used up in ohmic drop only. (Consequently there is no reactive voltage due to partial flux linkages within this type of conductor.)

The current density in the conducting annulus is

$$J=-\sigma 2B_s\frac{dr'}{dt}\text{ where }(r'<r\leq R) \tag{27}$$

The total current being carried by the wire is

$$i=\pi J(R^2-r'^2) \tag{28}$$

or

$$i(t)=-2B_s\pi\sigma(R^2-r'^2)\frac{dr'}{dt} \tag{29}$$

which can be integrated to find the coordinate r' as a function of time

$$\int_R^{r'}-2\pi B_s\sigma(R^2-r'^2)dr'=\int_0^t i(t)dt \tag{30}$$

The situation shown in Fig. 6 calls for $i(t)$ to be positive, since the induction is being changed to $+B_s$. If, for example, $i(t)$ in a sinusoidal function, then equation 30 would become $(0\leq\omega t\leq\pi)$

$$\frac{2}{3}\pi B_s\sigma[r'^3-3R^2r'+2R^3]=\frac{I_{\max}}{\omega}(1-\cos\omega t) \tag{31}$$

Three situations can be recognized in equation 31. First, the maximum current might not be sufficient to cause the surface of separation to move all the way to the center. An equivalent effect is caused by an increase in frequency. Under such conditions, the depth of penetration will be

$$\delta=R-r'_{\min}$$
$$\delta<R \tag{32}$$

which is computed from equation 13 as

$$\frac{2}{3}\pi B_s\sigma\delta^2(3R-\delta)=\frac{2I_{\max}}{\omega} \tag{33}$$

If now there is pronounced skin effect, i.e., $\delta<<R$, the depth of penetration is

$$\delta\cong\sqrt{\frac{I_{\max}}{\pi RB_s\omega\sigma}} \tag{34}$$

which expressed in terms of the maximum apparent surface current density, $J_{s\max}=(I_{\max}/2\pi R)$ amperes per meter, is

$$\delta\cong\sqrt{\frac{2J_{s\max}}{\omega\sigma B_s}} \tag{35}$$

The second situation which may be identified in equation 44 is that the current and frequency might be such that the surface of separation just reaches the center $(r'=0)$ at the end of a half-cycle. The current necessary to cause this condition is found from equation 31

$$I_{\max}=\frac{2}{3}\pi\omega\sigma B_sR^3\text{ amperes} \tag{36}$$

The third situation occurs if the current is so large that r' becomes zero before the end of the half-cycle; if so, the wire exhibits a constant resistance for the remainder of the half-cycle, equal to its d-c resistance.

The effective resistance may be computed for the case of pronounced skin effect $(\delta<<R)$. The loss density as a function of time will be in general, from equation 28

$$p(t)=\frac{J^2}{\sigma}=\frac{i^2(t)}{\pi^2\sigma(R^2-r'^2)^2}\text{ watts per meter}^3 \tag{37}$$

and the loss per meter of length will be

$$P(t)=\frac{i^2(t)}{\pi\sigma(R^2-r'^2)}\text{ watts per meter} \tag{38}$$

in which r' is a function of time which can be deduced from equation 30. The result is

$$P(t)\cong\sqrt{\frac{B_s}{2\pi R\sigma}}\,\frac{i^2(t)}{\sqrt{\int_0^t i(t)dt}}\text{ watts per meter} \tag{39}$$

If the current is a sinusoidal function

$$P(t)\cong I_{\max}{}^{3/2}\sqrt{\frac{B_s\omega}{2\pi R\sigma}}\left(\frac{\sin^2\omega t}{\sqrt{1-\cos\omega t}}\right)\text{watts per meter} \tag{40}$$

The average power dissipated in heat over the half-cycle will be

$$P_{av}\cong I_{\max}{}^{3/2}\sqrt{\frac{B_s\omega}{2\pi R\sigma}}\,\frac{1}{\pi}\int_0^\pi\frac{\sin^2\theta}{\sqrt{1-\cos\theta}}d\theta \tag{41}$$

or

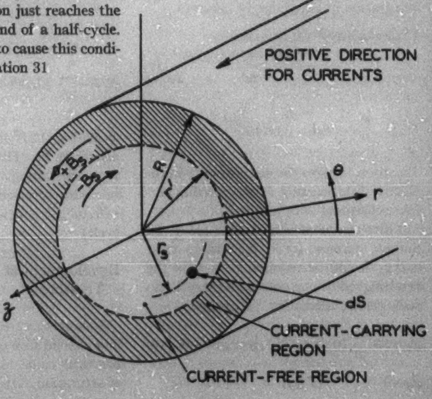


Fig. 6. Definition of symbols used in the cylindrical configuration

$$P_{av} \cong I_{max}^{3/2} \sqrt{\frac{B_s \omega}{2\pi R\sigma}} \frac{4\sqrt{2}}{3\pi} \text{ watts per meter} \quad (42)$$

When this power is expressed in terms of the depth of penetration, as computed from equation 34, it becomes

$$P_{av} \cong I_{max}^{2} \frac{8}{3\pi} \frac{1}{2\pi R\delta\sigma} \text{ watts per meter} \quad (43)$$

The effective resistance of the iron wire is defined by the expression $P_{av} = I_{rms}^2 R_{eff}$, yielding

$$R_{eff} \cong \frac{16}{3\pi}\left(\frac{1}{2\pi R\delta\sigma}\right) \text{ ohms per meter} \quad (44)$$

A conductor whose cross section is not circular, such as a rail, can be treated according to equation 44 if the perimeter of the cross section is substituted for the quantity $2\pi R$.

The depth of penetration, equation 35, is identical with that given by Rosenberg. It must be recognized, however, that there is a basic difference between the field models assumed by Rosenberg, Wolman, and others. This difference leads to higher losses computed according to the new treatment, as will be shown in equation 46. These similarities and differences will now be examined.

Rosenberg assumes that the current density, at any instant, is a maximum at the surface of the conductor and decreases at a uniform rate with depth, becoming zero at depth δ. Further, the current density at any level is in time phase with the current density everywhere else. Thus the depth of penetration and the depthwise average current density are related according to

$$I_{max} = 2\pi R\delta J_{max} \quad (45)$$

while the actual maximum current density, at the surface of the conductor, is $2J_{max}$.

According to the new treatment, there is an annular region at any instant in which the current density is uniform while the rest of the conductor carries no current. The thickness of this annulus varies with time and the current density within the annulus varies with time also; see equations 30 and 27. However, when the annulus is thickest, the current is maximum, and the current density at that instant is the same as the average given by Rosenberg, equation 45. Thus it is expected that both theories should give the same depth of penetration.

The ratio of loss per meter computed by Rosenberg to that computed according to the limiting treatment is

$$\frac{P_R}{P_L} = \left(\frac{2}{3}\frac{I_{max}^2}{2\pi R\delta\sigma}\right) \Big/ \left(\frac{8}{3\pi}\frac{I_{max}^2}{2\pi R\delta\sigma}\right) \quad (46)$$

or, the new treatment yields $(4/\pi) = 1.273$ times as much loss as Rosenberg's theory in the case of iron conductors with pronounced skin effect. The increased loss is attributed to the combined effect of phase shift within the field and the presence of time harmonics, both of which were neglected by Rosenberg.

Inductive Heating of Iron in Single-Phase Fields

The relationship between the surface field intensity and the total eddy current given by equation 20 for a plane configuration may be used for the treatment of inductive heating effects, if the least dimension of the surface to be heated is much greater than the depth of penetration. E.g., if the applied field intensity is a sinusoidal function, equation 20 becomes

$$2B_s\sigma z' \frac{dz'}{dt} = J_{smax} \sin \omega t \quad (47)$$

for each half-cycle. The location of the surface of separation between states $-B_s$ and $+B_s$ is

$$z' = \sqrt{\frac{1}{B_s\sigma}\int_0^t J_{smax} \sin \omega t dt}$$

where

$$\left(0 \le t \le \frac{\pi}{\omega}\right) \quad (48)$$

and the depth of penetration is

$$z_{max}' = \delta = \sqrt{\frac{2J_{smax}}{B_s\omega\sigma}} \text{ meters} \quad (49)$$

which is equivalent to equation 35.

The current density in the conducting layer will be

$$J = \frac{J_s}{z'} = \sqrt{B_s\omega\sigma J_{smax}} \frac{\sin \omega t}{\sqrt{1-\cos \omega t}} \text{ amperes per meter}^2 \quad (50)$$

and the loss per square meter of surface area is

$$P = z' \frac{J^2}{\sigma} = J_{smax}^{3/2} \sqrt{\frac{B_s\omega}{\sigma}} \left(\frac{\sin^2 \omega t}{\sqrt{1-\cos \omega t}}\right) \text{ watts per meter}^2 \quad (51)$$

The average loss is

$$P_{av} = \frac{4\sqrt{2}}{3\pi} J_{smax}^{3/2} \sqrt{\frac{B_s\omega}{\sigma}} \text{ watts per meter}^2 \quad (52)$$

which is the equivalent of equation 42. Thus for sinusoidal currents, the conductive and inductive heating processes are identical within the iron, and the same loss formula applies. This fact has been noted by Thornton in section 3.1 of reference 12.

Practical design of an inductive heating installation requires a means to compute the terminal voltage of the exciting winding. The real part of this voltage (in phase with the exciting current) will supply the resistive drop in the exciting circuit plus the equivalent resistance drop induced by the eddy-current losses. The reactive part will consist of the reactive drop induced by the eddy currents plus the reactive drop due to stray magnetic fields (e.g., the field existing in the space between the exciting winding and the surface to be heated). The real and reactive parts of the electric field induced by eddy currents will be computed next. The other components of terminal voltage may be computed by well-known methods.

The electric field induced in the exciting current sheet due to flux within the iron is given by equation 19, which when combined with equation 48 yields

$$E = \sqrt{\frac{B_s\omega J_{smax}}{\sigma}} \left(\frac{\sin \omega t}{\sqrt{1-\cos \omega t}}\right)$$

where

$$(0 \le t \le \pi/\omega) \quad (53)$$

which reduces to

$$E = \sqrt{\frac{2B_s\omega J_{smax}}{\sigma}} \left(\cos \frac{\omega t}{2}\right) \text{ volts per meter} \quad (54)$$

By Fourier analysis, the fundamental component is found to be

$$E_{fund} = \sqrt{\frac{2B_s\omega J_{smax}}{\sigma}} \left[\frac{8}{3\pi} \sin \omega t + \frac{4}{3\pi} \cos \omega t\right] \text{ volts per meter} \quad (55)$$

The phase angle between the electric field and the exciting current is

$$\theta = \tan^{-1} 0.5 = 26.6 \text{ degrees}$$

$$\cos \theta = 2/\sqrt{5} = 0.895 \quad (56)$$

Thus the new theory provides a means for computing the power factor of the load reflected into the exciting winding by the eddy currents. Such a computation was not possible using the Rosenberg treatment, since the effects of phase shifts within the field were not taken into account. The power factor of the eddy-current load is seen to be quite high. The relatively low power factor encountered in inductive heating installations is due to the reactive voltage of the unavoidably large stray magnetic fields.

It should be noted at this point that

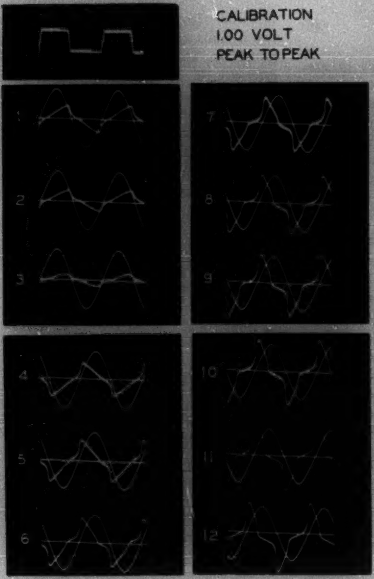


Fig. 7 (left). Induced electromotive force of a search coil wound on a solid iron ring excited with sinusoidal current

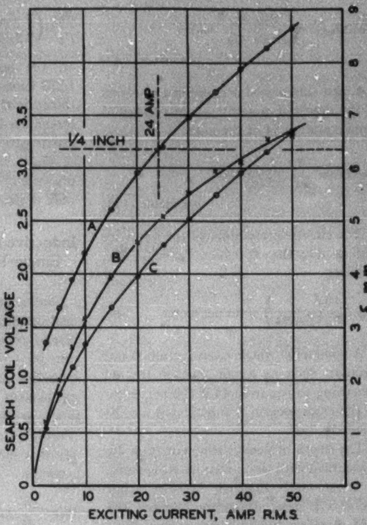


Fig. 8 (right). Computed and measured search coil voltage

A. Computed depth of penetration
B. Measured search-coil voltage
C. Computed search-coil voltage

the conductivity of iron varies considerably with temperature, and this effect should be taken into account when assigning values to σ in the various expressions. Of course the theory collapses entirely when the Curie point is reached; the magnetic behavior at this and higher temperatures follows the linear treatment with permeability equal to that of air.

Results of Experiment

The wave form of induced electromotive force as given by equation 54 has been checked experimentally, with the results being presented in Fig. 7 and Table I. Sinusoidal current was forced by connecting an adequate resistance in series with the exciting coil. The sinusoidal current wave is included for phase comparison, and its calibration varies among the different oscillograms. The voltage calibration is the same for all.

Table I. Key to Oscillograms of Fig. 7

Oscillogram	Exciting Current, RMS	Search Coil Voltage, Peak-to-Peak Volts
1	2.5	0.61
2	5.0	1.00
3	7.5	1.31
4	10.0	1.58
5	15.0	1.97
6	20.0	2.30
7	25.0	2.54
8	30.0	2.76
9	35.0	2.97
10	40.0	3.06
11	45.0	3.27
12	50.0	3.30

The specimen tested was a closed ring of square cross section, 0.5 inch by 0.5 inch, with mean diameter 5.375 inches, made of a particular grade of cast iron having good uniformity and high carbon content. Accordingly, the saturation induction is quite low and the resistivity is high, being 0.68 microhm-meter. Magnetic data are given in Table II. The ring was covered with two layers of ordinary black friction tape and then closely wound with 143 turns of American Wire Gauge-10 enamel-covered copper wire. A search coil of 12 turns was placed outside the exciting winding. Tests were made at 60 cycles.

The oscillograms show the characteristic sharp rise in voltage at the beginning of each half-cycle, and gradual decay, as given by equation 54. Furthermore, they show the distinctive behavior expected when the depth of penetration δ becomes approximately equal to, or greater than, half the thickness of the ring. The induced electromotive force is sharply reduced to a small value and remains there until the magnetic process starts again at the beginning of the next half-cycle, the peak voltage at the beginning of the half-cycle being unaffected by the later complete magnetic reorientation of the ring.

The peak-to-peak search coil voltage has been computed according to equation 54, further modified by Rosenberg's suggestion for determining B_s as discussed in the following section. The depth of penetration δ has been computed according to equation 49 using the same modification. These results as well as the measured search coil voltage, are presented in Fig. 8. The oscillograms show the characteristics of overexcitation at about 20 amperes exciting current, while the depth of penetration becomes equal to half the thickness of the ring at 24 amperes. (It should be noted again here that all computations are based upon the assumption of an infinite plane configuration.)

Table II. Magnetization Curve of Cast Iron Ring, by Fluxmeter

H, Amperes per Meter	B, Weber per Square Meter
0	0
1,000	0.55
2,000	0.76
3,000	0.85
4,000	0.90
6,000	0.97
8,000	1.01
10,000	1.04
12,000	1.05

Additional Remarks

A very detailed set of experiments has been reported recently by Thornton,[12] in which the losses in mild steel pipe were determined as a function of frequency and of current. The objective of the experiments was to present an empirical method of design for conductive and inductive heating of iron vessels and pipes. It was found that the loss at constant frequency varied as the 1.57 power of the current, while the new theory predicts an exponent of 1.50.

A somewhat empirical way of determining B_s, originally proposed by Rosenberg, does in fact lead to a loss exponent for this type of steel which is as near 1.57 as the accuracy of available information allows. Rosenberg proposed that the value of B_s to be used in a given problem should correspond to the maximum current carried per unit periphery of the conductor. If the magnetization curve for hot-rolled steel sheets (structural or "mild" steel) as given in handbooks is used in this way to compute the loss exponent, excellent agreement with Thornton's value is obtained. However, the agreement extends only over Thornton's range of test information which begins at about 7 oersteds, already well into the region where the induction ceases to change rapidly with applied field.

This empirical method of assigning B_s may be justified somewhat by inspection of equation 42. It is seen that the loss varies as the square root of B_s. The induction increases only moderately with increasing applied field when appreciable saturation is present, and the theory does not apply in the case of small saturation. In view of the small changes in B_s to be expected, and the further appearance of B_s under the radical, the method for taking the variation of B_s into account is not critical. Thus, the correlation between B_s and field intensity as obtained from the saturation curve could be expected to yield reasonably good results. In substantiation, the mechanics of the limiting nonlinear theory shows that the maximum field intensity is just equal to the maximum amperes per unit of periphery, so that higher values of induction than that taken from the saturation curve would not be expected.

Conclusions

It has been shown that the theory of alternating fields in solid iron proposed by Wolman and others[7-9], is a significant advance in that the objections to Rosenberg's treatment are overcome. Furthermore, a logical transition from the classical theory to the new theory has been demonstrated. The losses computed according to this theory are somewhat higher than those predicted by Rosenberg, as would be expected since the new theory accounts for both time harmonics of current density and phase shifts in the magnetic field.

Equations have been presented which show the usefulness of the new theory as a design method for inductive and conductive heating installations.

References

1. Losses of Eddy Currents in Electric Brakes and Dynamos, **R. Rudenberg.** *Sammlung Elektrotechnischer Vorträge*, Stuttgart, Germany, vol. X, 1906, pp. 269–370.

2. Solid Iron Conductors and Eddy Current Brakes, **E. Rosenberg.** *Elektrotechnik und Maschinenbau*, Vienna, Austria, vol. 41, no. 49, 1923, pp. 701–17.

3. Alternating Fields and Eddy Currents in Solid Iron with High Magnetic Saturation, **J. B. Barth.** *Electrotechnick Tidsskrift*, Oslo, Norway, vol. 48, no. 7, 1935, pp. 85–99.

4. Rise of Flux Due to Impact Excitation: Retardation by Eddy Currents in Solid Parts, **R. Pohl.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 96, pt. 2, 1949, pp. 57–65.

5. Influence of Eddy Currents on the Process of Flux Establisment, **S. Ya Dunaevskii.** *Elektrichestvo*, Moscow, U. S. S. R., no. 2, Feb. 1951, pp. 55–63.

6. Mathematical Description of Core Losses, **J. W. Hale, F. R. Richardson.** *AIEE Transactions*, vol. 72, pt. 1, Sept. 1953, pp. 495–501.

7. On the Eddy Current Delay in Magnetic Switching Phenomena, **W. Wolman, H. Kaden.** *Zeitschrift für Technische Physik*, Leipzig, Germany, vol. 13, no. 7, 1932, pp. 330–35.

8. The Alternating Field in Saturated Solid Iron, **G. Haberland, F. Haberland.** *Archiv für Elektrotechnik*, Berlin, Germany, vol. 30, 1936, pp. 126–33.

9. Application of Thin Permalloy Tape in Wide-Band Telephone and Pulse Transformers, **A. G. Ganz.** *AIEE Transactions (Electrical Engineering)*, vol. 65, April 1946, pp. 177–83.

10. Comparison of Methods of Analysis of Magnetic Amplifiers, **L. A. Finzi, G. F. Pittman, Jr.** *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 8, Jan. 1953, pp. 144–57.

11. Eddy-Current Phenomena in Ferro-Magnetic Materials, **H. M. McConnell.** *Technical Report No. 14*, Office of Naval Research Contract *N7 ONR 30306* and *30308*, *Project No. 975-272* and *275*, Carnegie Institute of Technology, Pittsburgh, Pa., Aug. 1953.

12. Resistance Heating of Mild-Steel Containers at Power Frequencies, **C. A. M. Thornton.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 99, pt. 2, 1952, pp. 85–93.

◆

Discussion

John F. H. Douglas (Marquette University, Milwaukee, Wis.): Hitherto it has been considered that the penetration of flux into iron castings with a-c excitation is much restricted. It has been somewhat of a mystery how output coefficients for eddy-current brakes and dynamometers could be as large as they are. This new theory seems to shed light on this problem. It is to be hoped that the author will find a method of analysis based on a closer approximation to actual saturation curves. The assumed *B-H* curve in Fig. 4 and the test results in Fig. 8 certainly show impressive agreement with theory as proposed:

Gordon J. Watt (Sperry Gyroscope Company, Great Neck, N. Y.): This discussion is written to acknowledge Mr. McConnell's presentation, offer some modifications which may extend the concept he proposes, and refer to some work done at Cornell University which is pertinent to this paper. I shall outline a different concept which follows a similar path of reasoning but ends with the same set of conclusions he presents. With due respect to Mr. McConnell's references, it is clear that a great deal of his own thought has gone into his concept of a square wave distribution of magnetic induction with time and depth in a ferromagnetic work piece.

In particular, a comparison of Fig. 1 and Fig. 4 seems to indicate that his concept is valid. His means of test by forcing sinusoidal current through an iron core solenoid are simple and the oscillograms presented in Fig. 7 are necessary to verify the theory he presents. If experience is any teacher, there are few who will comment on his paper and say his concept is right or wrong. Since his method of calculation yields predictable results, the concept itself does not wholly determine the importance of this work in the field. If the resulting equations may be used for design criteria in either of the fields he mentions, they are quite valuable. I have had occasion to apply a similar concept, yielding the same design criteria, to the design of shell motors. This work was printed in a copyrighted technical publication of the Sperry Gyroscope Company early in 1952. Tests result checked quite well with the theory presented there.

My own fundamental work was done at Cornell University and culminated in a thesis[1] which is on file in the school library. Until I read Mr. McConnell's reference to Mr. Ganz's application to thin permalloy tape, I was not aware of the latter's work. And so at this time I wish to acknowledge Mr. Ganz's fundamental idea as well as to compliment Mr. McConnell on his application of it to the concept which he presents. The alternate concept which I shall now discuss was presented in the thesis and based on the concept of a moving barrier between saturated and nonsaturated regions in a ferromagnetic material. References to illustrations pertain only to Mr. McConnell's paper.

Referring to Figs. 2 and 3 and Mr. McConnell's statement "if *H* is zero, the flux density is not necessarily changing," it is clear that this is a peculiar material which is not commonly used for the examples he has illustrated, viz., heating of an iron wire and inductive heating of iron. Let us consider what happens in the case of another peculiar material which is not representative of common materials. In this material the flux density is defined at zero in the region of H equal to zero. To accommodate the total flux, in a case similar to that outlined by Mr. McConnell, the lower boundary will move back toward the surface instead of a new boundary starting from the surface each time the total flux begins to decrease. The flux density below the lower boundary reduces to zero as the boundary moves outward. Following this line of reasoning yields equations 19 and 20 for voltage and current without a coefficient of two. Following through with a parallel reasoning used to obtain equations 47 through 56, expressions for current voltage and power are obtained which yield the same conclusions McConnell has reached. The predicted wave shapes are the same as those shown in Fig. 7, and the plot of voltage and skin depth versus current is as shown in Fig. 8. Therefore neither concept is confirmed or denied by the data presented in the paper under discussion.

Now suppose that a typical material is a combination of the two peculiar materials discussed and that the flux density in the typical material collapses to some residual value when H is in the vicinity of zero. Then the inner boundary will move outward and during the same cycle the outer boundary will move inward to accommodate the total flux. Will such a combination of these extreme concepts produce an equivalent result, to either of the extremes, no matter in what proportions they are mixed or whether or not they occur concurrently or in succession? My present considerations lead me to believe that the only difference will be in the time and space distribution of the current density in the penetrated layer. The equivalent impedance, average power loss, and depth of penetration will not be affected. These remarks are offered for what value they may be to Mr. McConnell or others doing work in this field.

Not having followed recent work in the field of induction heating, I was interested to read Mr. McConnell's reference to Thornton's work (ref. 12 of the paper). As a set of additional remarks to this discussion, I should like to recommend a theoretical approach to the practical case of predicting losses and skin depth as a function of applied field. This approach was worked out in my thesis,[1] and was made to cover the normal range of field strengths used in induction heating practice. By a method patterned after the work of H. F. Storm,[2] the work piece is broken down into a finite number of layers of constant thickness. An effective permeability and conductivity are assigned to each layer as determined by the B-H curve for the material and the temperature distribution assumed. These empirical values are inserted into a finite series.

Applied to the case of a mild steel, at constant frequency, this method produced a calculated power loss and skin depth as a function of field strength to the 1.61 and 0.39 power respectively. For low and high field strengths, these powers approached 1.50 and 0.50, having reached the extreme values mentioned above at normal field strengths used in induction hardening. No complete tests were run to verify this theory, but for a given steel it produced values corresponding to those of Brown, Hoyler, and Bierwirth for low field strengths,[3] and for high field strengths to those determined by the method presented in this paper. Mr. Thornton's data provide a further indication that the theory may be useful.

REFERENCES

1. A MATHEMATICAL STUDY OF INDUCTION HEATING PRINCIPLES, G. J. Watt. *Master of Electrical Engineering Thesis*, Cornell University, Ithaca, N. Y., June 1949.

2. SURFACE HEATING BY INDUCTION, Herbert F. Storm. *AIEE Transactions (Electrical Engineering)*, vol. 63, Oct. 1944, pp. 749–55.

3. RADIO FREQUENCY HEATING, George G. Brown, Cyril N. Hoyler, Rudolph A. Bierwirth. D. Van Nostrand Company, Inc., New York, N. Y., 1947.

H. M. McConnell: A depth of penetration of several millimeters can occur in a material such as mild steel, at power frequencies. This fact was recognized by Rosenberg, and his design calculations and test results on eddy-current brakes agree rather well. At that early stage, however, the design procedures were necessarily not as rational as our present knowledge allows. I too hope that this paper may be of use in the design of eddy-current brakes. I wish to thank Professor Douglas for his comments in this regard.

Mr. Watt raises the question of a comparison of his own treatment, developed in his thesis, with the one proposed in this paper. I propose to compare Mr. Watt's suggested approach with the one presented in this paper, in three situations for which experimental evidence is readily obtained.

The first situation is the application of a sinusoidally varying exciting current, as has been discussed in the paper. The induced electromotive force of a search coil enclosing all the resulting flux is sketched in Fig. 9(A). The same quantity predicted according to Mr. Watt's theory is sketched in Fig. 9(B). The corresponding experimental check has been given in the paper, and is presented in Fig. 7. A similar result is derived by a simplified analysis in reference 8 of the paper.

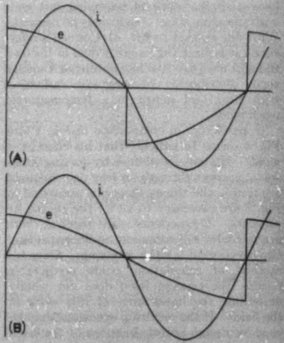


Fig. 9. Applying sinusoidally varying exciting current

A—From the paper
B—From Mr. Watt

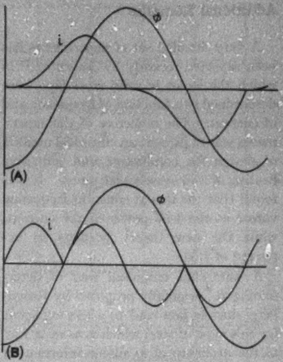


Fig. 10. Forcing sinusoidally varying flux to exist in the material

A—From the paper
B—From Mr. Watt

The second situation is the forcing of a sinusoidally varying flux to exist in the material. The exciting current for this condition is given in equation 22 and is shown in Fig. 10(A). The exciting current according to Mr. Watt's theory is shown in Fig. 10(B). This condition may be duplicated by experiment, approximately, if a sinusoidally varying voltage is applied to the terminals of the exciting winding on a closed solid iron core, and a sufficiently large capacitor is connected in parallel with the winding to offer a low impedance to harmonic currents. The results of such an experiment are displayed in Fig. 11. The sinusoidal trace is the voltage applied to the exciting winding, and the second trace is in

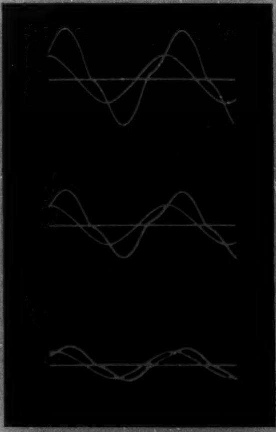


Fig. 11. Applying sinusoidally varying voltage to terminals of exciting winding

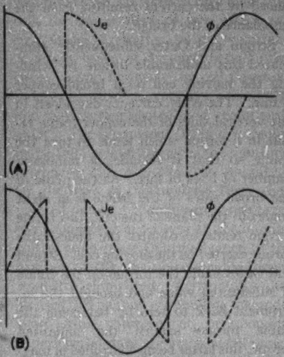


Fig. 12. Determination of wave form of current density

A—From the paper
B—From Mr. Watt

each case the resulting current flowing in the exciting winding. The rms exciting field intensities are approximately 5,000, 3,300, and 1,650 amperes per meter at 60 cycles. The winding and ring are the same as those used in the experiments described in the paper.

Fig. 11 shows the presence of both linear and saturated magnetization processes, as might be expected. Comparison of the theory given in the paper with that given by Mr. Watt may be made by observing the harmonic content of the exciting current.

Finally, a third situation investigated by Brailsford[1] is the determination of the wave form of the current density at various levels within a sheet of steel, when a sinusoidally varying total flux was caused to flow. This situation may be attacked by both the theory of this paper and Mr. Watt's theory, with the results presented in Figs. 12(A) and (B) for one such level. The results obtained experimentally by Brailsford agree well with Fig. 12(A).

I do not agree that Mr. Watt's equations and those developed in this paper yield the same losses. For example, in the case of sinusoidal flux, the losses predicted by Mr. Watt are 0.848 times those predicted in this paper. It is true that the dimensional behavior of the losses with respect to frequency and applied voltage are the same in the two theories. However, this agreement in dimensional behavior is due to our common assumption that a "boundary layer" exists. The other comparisons presented in this discussion prove that the basic mechanics of Mr. Watt's field model are not realistic.

A further demonstration of fallacy in Mr. Watt's field model may be obtained from the energy principle. Mr. Watt postulates that the "boundary layer" returns from the interior of the material to the surface during the second half-cycle in the case of sinusoidal current, or the second quarter-cycle in the case of sinusoidal flux. This statement can be supported only by the existence of a stored energy within the material. Let this stored energy accumulated over one interval be S; the material is relaxed at the beginning of a double interval. Also Mr. Watt's theory yields the result that the energy input from the excitation source is the same during each interval, half- or quarter-cycle, as the case may be. Let this input per interval be W. Let L be the energy loss during each interval; inspection of Mr. Watt's theory shows that L is the same for each. Therefore, the interval while the material is being magnetized yields $W=S+L$ and the next interval yields $W+S=L$. A solution is possible only if $S=0$, or the boundary layer cannot return toward the surface.

The process of energy flow demonstrated in this paper differs fundamentally from that assumed by Mr. Watt, as it is shown in the paper that energy always enters the material from its surface, propagating in waves toward the deepest penetration without reflection. This flow is characteristic of electromagnetic processes when no storage of energy occurs.

The basic differences between the two analyses pointed out in this closure give evidence of the dangers inherent to any reasoning based on extremely idealized models. It must be thoroughly understood in advance that even the limiting case of a phenomenon must obey all the physical laws which pertain to the phenomenon itself.

REFERENCE

1. INVESTIGATION OF THE EDDY-CURRENT ANOMALY IN ELECTRICAL SHEET STEELS, F. Brailsford. *Journal*, Institution of Electrical Engineers, London, England, vol. 95, pt. II, 1948, pp. 38–48.

Basic Theory and Experimental Verification of the A-C Galvanometer

THOMAS J. HIGGINS
MEMBER AIEE

WILLIAM KNEEN
NONMEMBER AIEE

Purpose and Scope

THE purpose of this paper is twofold: to advance an integrated account of the basic theory underlying the performance of the iron-cored a-c galvanometer when used as a null detector in an a-c bridge and to advance experimental verification of this theory.

Essentially, an a-c galvanometer comprises a moving coil—rectangular in shape, light in weight, and comprised of a number of turns of fine wire—vertically suspended between the poles of an electromagnet by two taut wires. The lower wire is anchored; the upper wire can be torsed by rotation of an externally mounted knob, thus enabling mechanical adjustment of the initial rest position of the moving coil and of the pointer fixed to it. The electromagnet comprises a laminated silicon-steel yoke excited by a winding supplied at a low voltage from the drop across one portion of a suitably subdivided resistor connected across a 110-volt a-c source. The magnetic field produced by the winding can be slightly modified by superposition of the eddy-current field induced in a steel strip which can be rotated at will over a limited angular range of position.

The circuit diagram pertinent to use of the a-c galvanometer as a null detector in an a-c bridge is typified in Fig. 1. Obviously, as the detecting element of an a-c bridge, the a-c galvanometer is connected in the bridge in the same manner as a D'Arsonval galvanometer in a d-c Wheatstone bridge. However, whereas the current through the moving coil of a D'Arsonval galvanometer has but one component (that stemming from the potential difference between the two junctions of the bridge to which its moving coil is connected), the current through the moving coil of the a-c galvanometer has two components: one stemming from the potential difference between the two junctions of the bridge to which the moving coil is connected, the other from the voltage generated in the moving coil by the alternating magnetic field of the electromagnet. In consequence, the net electromagnetic torque acting on the moving coil also comprises two components, one being produced by each of the two components of current. The component of torque stemming from the induced

---

Paper 54-171, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted October 19, 1953; made available for printing December 31, 1953.

THOMAS J. HIGGINS is with the University of Wisconsin, Madison, Wis., and WILLIAM KNEEN is with The Consolidated Engineering Corporation, Pasadena, Calif.

This paper is based upon a thesis supervised by Professor Higgins and submitted by Mr. Kneen in partial fulfillment of the requirements for the degree of Master of Science in electrical engineering.

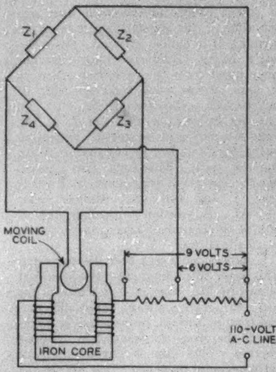


Fig. 1. Circuit diagram of galvanometer connected as a null detector in an a-c bridge

voltage is, in general, much larger than the component of torque stemming from the unbalance of the bridge. In consequence, use of the galvanometer as a null detector for the component of current stemming from the bridge circuit is possible only if the component of torque due to the induced component of current is eliminated or otherwise manipulated so that the galvanometer coil is sufficiently insensitive to it.

Because of a sensitivity comparable with that of the D'Arsonval galvanometer, those books on electrical measurements which mention the iron-cored a-c galvanometer cite it as that moving coil instrument best suited to use as a null detector for power-frequency a-c bridges. Yet, of the mere six of the senior author's collection of all consequential books in English, French, and German on electrical measurements which do mention and so cite the a-c galvanometer, only two state the afore-mentioned difficulty which is encountered in trying to use the instrument as a null detector, and none discuss means of overcoming this difficulty. Further, none of these books advance the analytic theory of the instrument wherefrom a user might deduce means of eliminating the undesirable torque due to the induced component of current. Accordingly, the prime purposes of the present paper are to fill in these gaps and to advance specific information enabling ready use of the a-c galvanometer.

The essential content of the paper may be epitomized as follows:

1. References 1 through 9 are summarized for consequential content. This review evidences that Terry[9] first suggested that the undesirable torque resulting from the current induced in the moving coil of the galvanometer can be eliminated by inserting capacitive or inductive (as necessary) reactance in series with the moving coil, of such magnitude as to render the equivalent impedance of the circuit containing the moving coil a pure resistance.

2. An analytic equation for the performance of the a-c galvanometer when used as a null detector in an a-c bridge is deduced[7] from Weibel's general mathematical theory of the response of the a-c galvanometer to an impressed periodical electromotive force of arbitrary wave shape. This equation yields theoretical confirmation of the soundness of Terry's suggestion.

3. Account is given of the experimental determination of the magnitude of the series capacitance required to eliminate the undesired torque when the a-c galvanometer is used as a null detector in a Maxwell inductance bridge, and of the magnitude of series inductance required when used as a detector in a Wien capacitance bridge. The experimental values prove to be in excellent agreement with corresponding theoretically calculated values. This excellent agreement of experimental and theoretical values substantiates the correctness of the analysis remarked in item 2 and indicates both that the method of eliminating the undesired torque by insertion of a proper magnitude of reactance in series with the moving coil is satisfactory in practice and that the correct value of reactance therefore can be easily calculated.

Review of Previous Work

The first account of the use of the iron-cored a-c galvanometer as a null detector in an a-c bridge appears to be that by Stroud and Oates.[1] They conjectured that inasmuch as this galvanometer is capable of great sensitivity it ought to prove an ideal detector for low-frequency a-c bridges as used for the measurement of capacitance or inductance. In confirmation of this conjecture they built and tested several instruments.

They wound the moving coil of their first instruments on a wooden frame, but this gave rise to the following difficulty. By well-known theory, if a closed coil is suspended in an alternating magnetic field, the current induced in the coil results in an electromagnetic torque that causes the coil to take up a position such that it includes a minimum number of lines of flux and also causes it to be very stable in this position. The moving coil of an a-c galvanometer is always closed through external impedance when used as a null detector in an a-c bridge. In consequence, the voltage generated in the moving coil by the alternating magnetic field of the electromagnet produces a current which results in a strong torque which, when the bridge is near balance, obscures the much weaker torque produced by the current resulting from the unbalance of the bridge.

Stroud and Oates subsequently eliminated this undesirable torque by winding the moving coil on a broad copper frame. The eddy currents developed in the vertical sides of the copper frame result in a torque which tends to turn the frame so that it includes a minimum number of lines of flux. As the plane of the broad sides of the frame is at right angles to the plane of the coil, this frame torque tends to counter the undesirable torque exerted on the moving coil. Stroud and Oates made the torque on the frame of substantial value and brought the instrument back to zero by balancing the frame torque against the undesired torque, this latter being controlled in magnitude by shunting the moving coil with a resistor.

In 1905 Terry published an article which encompassed, more or less, a continuation of the work initiated by Stroud and Oates. Terry's article is primarily concerned with the measurement of the temperature coefficient of capacitors (Stroud and Oates had earlier conjectured that capacitors have a temperature coefficient) through the use of an a-c galvanometer. He, more specifically than Stroud and Oates, pointed out the prime difficulty to be encountered in the use of this instrument stating "The great difficulty in manipulating the instrument is due to the fact that a closed coil suspended in an alternating magnetic field takes up a position of great stability such that its axis is at right angles to the field or so that the coil itself will include a minimum number of lines of force. Since, in any use of the instrument the coil must necessarily be closed through a greater or less resistance, this stability reduces the sensibility enormously."

Like Stroud and Oates, Terry found that the use of a copper frame necessarily limited the range of sensitivity of the instrument. Thus, if there was sufficient capacitance in the external circuit, it would produce a leading current in the moving coil. In turn, this current would produce a torque which would have the same effect as the torque on the sides of the frame, making it necessary to shunt the moving coil with such a small resistor that sensitivity was greatly reduced.

The essential origin of the undesired torque is easily noted by the aid of the vector diagram of Fig. 2. Thus, quoting Terry's article, "If **OA** represents the flux in the field, then the vector **OB** would represent the current induced by it in the galvanometer coil, providing it had no inductance but simply resistance. There would then be no force acting upon the coil, for the product of two vectors at right angles is zero (i.e., torque = current × flux × cosine of the angle between the two). Since, however, the coil possesses inductance, the current lags behind **OB** and may be represented by a vector such as **OC**. The component **OD** of this vector being 180° behind **OA** gives the repulsion producing the objectionable stability." On the basis of this analysis Terry then pointed out that if sufficient capacitance were connected in series with the moving coil so as to cause **OC** to advance to the position of **OB**, the objectionable torque would be eliminated. Terry states that he used this method of annulling the undesired torque; however, no corresponding experimental data are given in his paper. In virtue of the resulting sensitivity of the a-c galvanometer when so compensated, Terry was able to establish that capacitances had an appreciable negative temperature coefficient.

In 1906 Abraham[3] pointed out that, if there is any dissymmetry in the construction of an a-c galvanometer, the alternating field generated by eddy currents in the fixed metal parts will produce a strong turning moment on the moving coil, in one direction or the other. This fact suggests that the objectionable torque due to the induced current might be eliminated by deliberately introducing a disturbing metallic element which will result in a torque which is equal and opposite to the undesired torque. (The instrument used in the experimental work to be described encompasses such an element, a bent steel strip, for overcoming the objectionable torque; however, this device proved inadequate when the galvanometer was used in highly reactive circuits.) Abraham also pointed out that a series capacitor (as suggested by Terry) can be used to annul the undesired torque produced by the induced current in the moving coil.

In 1906 Trowbridge and Taylor[4] discussed the use of the a-c galvanometer for the measurement, by comparison, of capacitances and inductances. By use of a special bridge circuit nearly free from distributed capacitances they were able to make measurements to an accuracy of one part in 10,000.

In 1907 Taylor[5] rediscussed much of what Terry had covered earlier. In particular, he stressed that the direction of the undesirable torque produced on the moving coil depends on whether the current leads or lags the induced voltage producing it.

In 1908 Sumpner and Phillips[6] discussed some of the various uses for which they had employed the a-c galvanometer. In designing an instrument for these uses, they made a special effort to obtain the magnetic field as nearly in quadrature with the exciting voltage as was possible, thus attempting to annul the undesired torque through obtaining a 90-degree phase angle between the flux of the main field and the current resulting from the induced voltage in the moving coil.

In 1918 Weibel[7] carried out a thorough mathematical analysis of the motion of the moving coil of an a-c galvanometer in response to a periodic impressed voltage of arbitrary wave form. He also analyzed the important operating constants of the instrument, their effect on the motion of the coil, and the effects of circuit reactance on the action of the instrument.

Recently, Dalmijn[8] and Milatz, Endt, Alkemade, and Olink[9] have given an account of certain interesting uses and properties of the a-c galvanometer, but this work has no direct bearing on the prime interest of this paper.

Fundamental Theory of Operation as a Null Detector

In the preceding section it is remarked that if a closed coil is suspended in an alternating magnetic field the torque resulting from the induced current will cause it to take a position such that it includes a minimum number of lines of flux and it will be very stable in this position. Terry suggested that if the impedance in series with the moving coil is made purely resistive this undesired torque due to the induced current will be annulled. Analytically, this fact can easily be deduced from Weibel's analysis. Thus, the governing equation of motion as deduced by Weibel is

$$K\ddot{\theta}+\left(D+\sum_{n=1}^{\infty}\frac{G_n^2}{Z_n}\cos\gamma_n\right)\dot{\theta}+\left(U-\sum_{n=1}^{\infty}\frac{n\omega G_n^2}{Z_n}\sin\gamma_n\right)\theta=\sum_{n=1}^{\infty}\frac{G_nE_n}{Z_n}\cos(\alpha_n-\beta_n-\gamma_n)-\theta_0\sum_{n=1}^{\infty}\frac{n\omega G_n^2}{Z_n}\sin\gamma_n \quad (1)$$

where

K = the moment of inertia of the moving system
θ = the angular deflection from the reference position
D = damping constant
G_n = the rate of change with deflection of the effective value of the nth harmonic of the total flux
Z_n = the impedance of the complete moving coil circuit to the current of the nth harmonic frequency
γ_n = the phase angle of lead of the nth harmonic of the current ahead of the electromotive force (emf) producing it.
U = the moment of restoration acting on the moving system
n = the degree of harmonic ($n=1$ for the fundamental frequency)
$\omega=2\pi$ times the fundamental frequency
E_n = the effective value of the nth harmonic of the impressed emf
α_n = the phase angle of lead of the nth harmonic flux ahead of the fundamental flux
β_n = the phase angle of lead of the nth harmonic of the impressed emf ahead of the fundamental flux
θ_0 = the deflection for no flux through the coil

For the case of a sinusoidal magnetic field and a sinusoidal impressed emf of the same frequency, equation 1 reduces to

$$K\ddot{\theta}+\left(D+\frac{G^2}{Z}\cos\gamma\right)\dot{\theta}+\left(U-\frac{\omega G^2}{Z}\sin\gamma\right)\theta=\frac{GE}{Z}\cos(\beta+\gamma)-\theta_0\frac{\omega G^2}{Z}\sin\gamma \quad (2)$$

Under steady-state conditions, $\ddot{\theta}$ and $\dot{\theta}$ are each zero and equation 2 yields

$$\theta_{ss}\left[U-\frac{\omega G^2}{Z}\sin\gamma\right]=\frac{GE}{Z}\cos(\beta+\gamma)-\theta_0\frac{\omega G^2}{Z}\sin\gamma \quad (3)$$

If the equivalent impedance in series with the moving coil is changed by adding impedance so as to finally obtain a pure resistance circuit, γ then equals zero and equation 3 gives

$$\theta_{ss}U=\frac{GE}{Z}\cos\beta \quad (4)$$

and thus

$$\theta_{ss}=\frac{GE}{UZ}\cos\beta \quad (5)$$

Inasmuch as $Z=R'$, equation 5 can be written as

$$\theta_{ss}=\frac{GE}{R'U}\cos\beta \quad (6)$$

We note that equation 6 indicates that

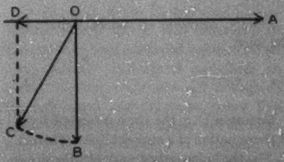


Fig. 2. Vector diagram of a-c galvanometer

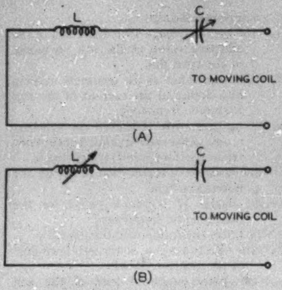


Fig. 3. Circuits for checking condition of zero torque in pure resistance circuit

the steady-state deflection is independent of the induced voltage in the moving coil when the equivalent impedance of the circuit containing the moving coil is purely resistive.

Deflection for No Flux Through Coil

Under steady-state conditions equation 2 reduces to

$$\left(U-\frac{\omega G^2}{Z}\sin\gamma\right)\theta_{ss}=\frac{GE}{Z}\cos(\beta+\gamma)-\theta_0\frac{\omega G^2}{Z}\sin\gamma \qquad (7)$$

If no emf is impressed in the circuit of the moving coil, $E=0$, and equation 7 gives

$$\left(U-\frac{\omega G^2}{Z}\sin\gamma\right)\theta_{ss}=-\theta_0\frac{\omega G^2}{Z}\sin\gamma \qquad (8)$$

Solving for θ_{ss} yields

$$\theta_{ss}=\frac{\theta_0}{1-\dfrac{UZ}{\omega G^2\sin\gamma}} \qquad (9)$$

Equation 9 can be written as

$$\theta_{ss}=\frac{\theta_0}{1-kx} \qquad (10)$$

where

$$x=\frac{Z}{\sin\gamma} \qquad (11)$$

$$k=\frac{U}{\omega G^2} \qquad (12)$$

Finally, equation 10 can be written in the form

$$x=-\frac{\theta_0}{k}\theta_{ss}^{-1}+\frac{1}{k} \qquad (13)$$

Equation 13 is the slope intercept form of the equation of a straight line. Therein θ_{ss}^{-1} is the independent variable, the slope is $-\theta_0/k$, and the ordinate intercept is $+1/k$. Obviously, if $x=Z/\sin\gamma$ is plotted as a function of θ_{ss}^{-1}, the slope and the intercept of the resulting curve can be measured and θ_0 determined therefrom, since

$$(-\text{slope/ordinate intercept})=\theta_0 \qquad (14)$$

This analysis suggests a means of experimentally checking Weibel's theory. Thus, the data necessary to plotting the curve of equation 13 can be obtained by connecting the moving coil to a series combination of resistor and capacitor, the resistor set at a convenient fixed value which serves to properly limit the maximum deflection, and the capacitor then varied to obtain the desired data. In obtaining such data the following is to be kept in mind. Since the pointer position corresponding to $\theta_0=0$, or $\theta_{ss}=0$ (when the initial position is at $\theta_0=0$) is unknown at the outset, it is logical to set the instrument mechanically to the scale zero, but $\theta_0=0$ may not be coincident with this position. The resulting value from equation 14 for θ_0 will therefore be the number of divisions to the right or left of the scale zero that locates $\theta_0=0$. The reason for such is, of course, that the data for the determination of θ_0 are taken with the instrument initially at scale zero rather than $\theta_0=0$. As found in the experimental investigation mentioned in the section entitled "Experimental Confirmation of Theory," the difference between $\theta_0=0$ and scale zero proved to be very small for the instruments used in the investigation.

Zero Deflection as an Indication of Pure Resistance

When no emf is impressed in the external circuit, as in the case where a series combination of inductance and capacitance is connected directly across the terminals of the moving coil, the steady-state deflection is given by equation 9. Inasmuch as when $\gamma=0$, the denominator of equation 9 becomes infinite, and thus $\theta_{ss}=0$, equation 10 suggests a means of measuring inductance or capacitance by use of a calibrated variable capacitor or inductor in a circuit such as evidenced in Fig. 3.

Experimental Confirmation of Theory

Apparatus

The galvanometers used in the experimental work were two Leeds and Northrup a-c galvanometers, serial numbers *263,585* and *502,747*. The resistance of the moving coil of the former was 254 ohms; that of the latter was 20.5 ohms.

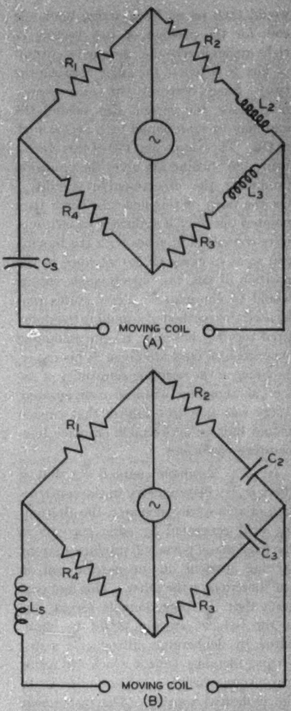


Fig. 4. Test circuits using (A) Maxwell inductance bridge, and (B) Wien capacitance bridge

The zero adjustments, bent steel strips, were removed from the galvanometers for the course of the experimental work. Additional pieces of equipment included several Cornell-Dubilier decade capacitor boxes, General Radio decade capacitor boxes, General Radio variable inductors, and General Radio fixed inductors.

Purpose

The experimental work fell into two parts. In the first part the galvanometers were used as detectors in bridge circuits. The purpose of this portion of the investigation was not to measure inductance or capacitance but to obtain verification of the theoretically indicated fact, that the undesired torque due to the induced current can be eliminated by connecting inductance or capacitance in series with the moving coil so that, effectively, the latter is connected in a pure resistance circuit. The second part of the

Table I. Data on the Inductance Bridge, Fig. 4(A)

R_1 and R_4, Ohms	R_2 and R_3, Ohms	L_s and L_3, Millihenrys (mh)	C_s Calc., Microfarads (Mfd)	C_s Exp., Mfd	ΔL or ΔR Units Indicated	Per-Cent Change in L or R	Number of Div. Change
10	429	500	28.16	36	on zero		
		500	28.16	36	L_s of 50 mh	10.0	7.0
		500	28.16	36	R_3 of 10 ohms	2.3	6.5
		500	28.16	46	L_s of 50 mh	10.0	4.0
		500	28.16	46	R_3 of 10 ohms	2.3	4.5
		500	28.16	30	L_s of 50 mh	10.0	12.0
		500	28.16	30	R_3 of 10 ohms	2.3	14.5
		400	35.1	47	on zero		
		400	35.1	47	L_s of 50 mh	12.5	5.2
		400	35.1	47	R_3 of 10 ohms	2.3	7.5
		400	35.1	60	L_s of 50 mh	12.5	4.0
		400	35.1	60	R_3 of 1 ohm*	0.23	4.8
		400	35.1	40	L_s of 50 mh	12.5	7.5
		400	35.1	40	R_3 of 10 ohms	2.3	13.0
		300	47.0	73	on zero		
		300	47.0	73	L_s of 50 mh	16.6	4.5
		300	47.0	73	R_3 of 10 ohms	2.3	7.5
		300	47.0	63	L_s of 50 mh	16.6	6.0
		300	47.0	63	R_3 of 10 ohms	2.3	10.0
		300	47.0	80	L_s of 50 mh	16.6	4.0
		300	47.0	80	R_3 of 10 ohms	2.3	6.5
	32	50†		∞	L_s of 5 mh	10.0	2.0
	32	50†		∞	R_3 of 1 ohm*	3.1	3.0

*Deflection too large with change of 10 ohms in *R*.

†When *L* becomes as low as 100 mh, capacitance is no longer needed to zero the galvanometer. For the above values $C_s = \infty$ (short-circuited).

Note: Data not taken with measurement of inductance as aim but to determine the effect of the compensating impedance on sensitivity. Galvanometer no. *592,747* was used.

Table II. Data on the Capacitance Bridge, Fig. 4(B)

R_1 and R_4, Ohms	R_2 and R_3, Ohms	C_2, Mfd	C_3, Mfd	L_s Calc., Henrys	L_s Exp., Henrys	Change in C from C_3, Mfd	Div. Defl.
1,000	10	2	1.985	1.76	1.76		
		2	1.985	1.76	1.76	0.001	2.5
		2	1.985	1.76	2.76	0.001	0.5
		2	1.985	1.76	$L_s < 1.76$*		
		1	0.955	3.52	3.52		
		1	0.955	3.52	3.52	0.001	2.0
		1	0.955	3.52	4.00	0.001	0.5
		1	0.955	3.52	$L_s < 3.52$*		
		3	2.942	1.172	1.172		
		3	2.942	1.172	1.172	0.001	4.0
		3	2.942	1.172	1.5	0.001	0.1
		3	2.942	1.172	$L_s < 1.172$*		
		4	4.197	0.879	0.879		
		4	4.197	0.879	0.879	0.001	0.7
		4	4.197	0.879	1.0	0.001	0.0
		4	4.197	0.879	$L_s < 0.879$*		
		5	4.942	0.704	0.704		
		5	4.942	0.704	0.704	0.001	1.5
		5	4.942	0.704	1.0	0.001	0.0
		5	4.942	0.704	$L_s < 0.704$*		
		6	5.759	0.586	0.586		
		6	5.759	0.586	0.586	0.001	2.0
		6	5.759	0.586	1.0	0.001	0.0
		6	5.759	0.586	$L_s < 0.586$*		

*When L_s is less than indicated value, it is not possible to zero the instrument.
Note: Galvanometer no. *502,747* was used.

investigation was devoted to experimental investigation of the theory, deduced from Weibel's equation of motion of the moving coil, outlined in the two preceding sections.

Data Taken and Associated Circuits

The data of Table I were taken with the Maxwell inductance bridge of Fig. 4(A). To obtain these data the bridge was balanced by setting its impedance values so as to satisfy the known equations of balance, and the series capacitance then was adjusted until the pointer stood at zero. The data of Table II were taken with the Wien capacitance bridge of Fig. 4(B). In obtaining these data the bridge was balanced by setting its impedance values so as to satisfy the known equations of balance, and the series inductor then was adjusted until the pointer stood at zero. This experimental work serves to check the efficacy of the method of eliminating the undesired torque originally suggested by Terry.

Data are also given in Tables I and II for values of deflection corresponding to values of *L* and *C* greater and smaller than the calculated value of the compensating impedance. It is to be noted that the sensitivity of the instrument is greatly affected by the value of the compensating impedance.

A series of experiments was then run to enable comparison of the sensitivities of the two galvanometers. The two instruments differ in the resistance of the moving coils: galvanometer no. *263,585* has a coil resistance of 254 ohms; galvanometer no. *502,747* has a coil resistance of 20.5 ohms. The results of these experiments are given in Tables III and IV.

Considerable data were then taken to enable investigation of the theory advanced in the two preceding sections. Limitations of space and time forestall detailed discussion of the experimental details and the analysis of the data thus obtained. The reader interested in a full account of this phase of the investigation is referred to the thesis (available on loan) on which this paper is based.[10] Experiment yielded values providing complete substantiation of the theory advanced in the two preceding sections. In corroboration of this, see Fig. 5, which evidences experimental confirmation of the straight-line relationship predicated in equation 13.

Interpretation of Experimental Data

INDUCTANCE BRIDGE

Examination of the data of Tables I and IV reveals that the galvanometer is most sensitive when the series compensating capacitance is as small as possible. In other words, with less capacitance than the value required to make the phase angle zero or less than enough to make it possible to bring the instrument to its mechanical zero point, a greater deflection may be obtained for a given change in inductance or resistance than when the value is greater. Thus it may be advisable to use a "false zero," since the sensitivity decreases very rapidly as the value of *C* is increased.

CAPACITANCE BRIDGE

In the case of the capacitance bridge, Tables II and III, it was necessary to have at least the calculated value of series compensating inductance required to make the phase angle zero, for as soon as the inductance was decreased below this value the instrument rapidly fell off from zero. The sensitivity to changes in capacitance apparently was best when the compensating inductance was the mini-

Table III. Data for Galvanometer Comparison, Capacitance Bridge, Fig. 4(B)

R_1 and R_4, Ohms	R_2 and R_3, Ohms	C_2 and C_3, Mfd	L_s Calc., Henrys	L_s Exp., Henrys	Change in R or C from R_4 or C_3	Div. Defl.	Galvanometer No.
10	100	2.04	1.76	1.76		*	*263,585*
				1.76	0.01 mfd	1.3	
				1.76	100 ohms	1.0	
				2.00	0.01 mfd	0.5	
				2.00	100 ohms	0.5	
				1.68	0.01 mfd	4.5	
				1.68	100 ohms	2.5	
10	100	2.04	1,76	1.76		0*	*502,747*
				1.76	0.01 mfd	2.0	
				1.76	100 ohms	2.0	
				2.00	0.01 mfd	1.3	
				2.00	100 ohms	2.2	
				1.68	0.01 mfd	2.3	
				1.68	100 ohms	1.5	

* Galvanometer no. *263,585* was still not on zero with 1.76 henrys compensating inductance and fell away from zero rapidly below this value. Galvanometer no. *502,747* was almost on zero with 1.35 henrys and on zero with 1.76 henrys compensating inductance.

Table IV. Data for Galvanometer Comparison, Inductance Bridge, Fig. 4(A)

R_1 and R_4, Ohms	R_2 and R_3, Ohms	L_2 and L_3, Mh	C_3 Calc., Mfd	C_3 Exp., Mfd	Change in R or L from R_4 or L_3	Div. Defl.	Galvanometer No.
10	437	500	28.16	28		*	*502,747*
				28	25 mh	14	
				28	4 ohms	13	
				35	25 mh	4	
				35	4 ohms	5	
				27	25 mh	17	
				27	2 ohms†	2	
10	437	500	28.16	28		*	*263,585*
				28	25 mh	4.5	
				28	4 ohms	3.5	
				35	25 mh	1	
				35	4 ohms	1	
				27	15 mh.†	8.5	
				27	2 ohms†	5	

*Galvanometers were not quite on zero with the calculated necessary 28 mfd, but 32 mfd zeroed the instruments.

†Smaller changes had to be used so the deflection remained on the instrument's scale.

Note: The stops on the swing of the needle of galvanometer no. *502,747* were at ±13 divisions. The stops on the swing of the needle of galvanometer no. *263,585* were at ±7 divisions.

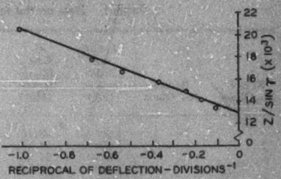


Fig. 5. Determination of $\theta_0=0$ referred to scale zero. Galvanometer no. 502,747. R = 2,000 ohms. C varied from 0 to 0.2 mfd. (From data of Table 6[10])

Slope = −7600 = $-\theta_0/k$
Intercept = 13000 = $1/k$
$\theta_0=0$ is +0.585 division to right of scale zero

mum necessary to keep the needle from going off scale (with the galvanometer of low coil resistance, Table III, the compensating inductance can be lowered somewhat below the calculated value). Increasing the inductance above the calculated value in every case reduced the sensitivity to changes in capacitance.

The data in Tables III and IV serve to show that the instrument with the lower resistance was most sensitive in the inductance bridge and somewhat less sensitive in the capacitance bridge circuit than the instrument with the higher resistance.

Conclusions

The previous work on theory and application of the iron-cored a-c galvanometer is reviewed, revealing that Terry[2] first suggested that the undesirable torque resulting from the current induced in the moving coil of the galvanometer could be eliminated by inserting reactance, capacitive or inductive as needed, in series with the moving coil. The magnitude of this reactance is to be such that the resulting equivalent impedance of the circuit containing the moving coil is pure resistance.

A theoretical confirmation of Terry's suggestion is deduced[7] from Weibel's general mathematical theory of the response of the galvanometer to an impressed periodical electromotive force of arbitrary wave shape.

Experimental determination of the magnitude of series capacitance required to eliminate the undesired torque when the galvanometer is used as a null detector in a balanced Maxwell inductance bridge is in agreement with the theoretically calculated value. Similarly, experimental determination of the magnitude of series inductance required to eliminate the undesired torque when the galvanometer is used as a null detector in a balanced Wien capacitance bridge is in agreement with the theoretically calculated values. This excellent agreement between experimental and theoretical values substantiates the correctness of the mathematical analysis mentioned in the preceding paragraph and indicates that the method of eliminating the undesired torque by means of inserting a proper magnitude of reactance in series with the moving coil is satisfactory in actual practice.

Experimental determination of the pointer position of the scale corresponding to $\theta_0=0$ is effected for several different values of equivalent resistance of the moving coil circuit. The effect of varying θ_0 is in agreement with the effect indicated by the mathematical theory of the behavior of the a-c galvanometer when used as a detector in an a-c bridge, as worked out in the section entitled "Deflection for No Flux Through Coil," on the basis of Weibel's general theory. This agreement between theoretical and experimental values provides additional confirmation of the correctness of the desired theory of behavior.

References

1. On the Application of A.C. to the Calibration of Capacitance Boxes, and to the Comparison of Capacitance and Inductance, **W. Stroud, J. H. Oates.** *Philosophical Magazine*, London, England, vol. 6, 1903, 6th series, pp. 707–20.

2. On the Variation of a Capacity with Temperature, **E. M. Terry.** *Physical Review*, New York, N. Y., vol. 21, 1905, pp. 193–97.

3. Galvanomètre à Cadre Mobile Pour Courants Alternatifs, **H. Abraham.** *Comptes Rendus*, Paris, France, vol. 142, 1906, pp. 993–94.

4. Note on the Comparison of Capacities, **A. Trowbridge, A. H. Taylor.** *Physical Review*, New York, N. Y., vol. 23, 1906, pp. 475–88.

5. Theory of Control of the Alternating Current Galvanometer, **A. H. Taylor.** *Physical Review*, New York, N. Y., vol. 25, 1907, pp. 61–63.

6. A Galvanometer for Alternate Current Circuits, **W. E. Sumpner, W. C. Phillips.** *Proceedings*, Physical Society, London, England, vol. 22, 1908, pp. 395–409.

7. A Study of Electromagnet Moving Coil

Galvanometers for Use in Alternating Current Measurements, E. Weibel. *Bulletin*, U. S. Bureau of Standards, Washington, D. C., vol. 14, 1918, pp. 23–58.

8. A. Dalmijn. *Electro-techniek*, The Hague, Netherlands, Aug. 14, 1947.

9. The Alternating-Current Galvanometer, M. W. Milatz, P. M. Endt, C. Alkemade, J. Olink. *Physica*, The Hague, Netherlands, vol. 14, 1948, pp. 260–68.

10. Theory of the Alternating-Current Galvanometer and Experimental Verification, W. Kneen. *M.S. Thesis*, Illinois Institute of Technology, Chicago, Ill, Jan. 1949.

◆

Discussion

Perry A. Borden (The Bristol Company, Waterbury, Conn.): Somewhat belatedly, this contribution to the art of a-c measurement comes as an answer to a problem mentioned in my 1923 paper.[1] In the closure of that paper, I referred to the futility of efforts to detect balance conditions by the use of an iron-cored a-c galvanometer. For that reason it was found necessary to develop and use the "asynchronous" method described in the body of the paper. Briefly, that method comprised impressing the unbalance potential upon one winding of an electrodynamic instrument whose other winding was excited from a source having a slightly different frequency. The resulting "beating" action could readily be observed, and a balance detected by its eventual reduction to a zero magnitude. As an alternative, the potential could be "rectified" by a commutating switch operated slightly away from the test frequency and the commutator output applied to a conventional d-c galvanometer. Again, the beating became apparent, and its reduction could be made a criterion of balance condition in the network.

Germane to the subject matter of the current paper is the work of Gilbert[2,5] and others[3–4] on galvanometers having a-c and d-c magnitudes superposed in a common deflecting coil.

References

1. Balance Methods in Alternating Current Measurement, Perry A. Borden. *AIEE Transactions*, vol. 42, 1923, pp. 395–405.

2. The Induction Galvanometer, a Sensitive Instrument Converter, R. W. Gilbert. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1121–26.

3. Galvanometer, P. A. Borden. U. S. Patent No. *2,368,701*, Jan. 6, 1945.

4. D. C. Current-Detector, J. R. Pattee. U. S. Patent No. *2,405,049*, July 30, 1946.

5. Measuring and Control Apparatus, R. W. Gilbert. U. S. Patent No. *2,486,641*, Nov. 1, 1949.

Thomas J. Higgins and **William Kneen**: In answer to a query raised in the discussion following the oral presentation of the paper we remarked on the ready availability on loan of Mr. Kneen's M. S. thesis to anyone desiring to read the details of the substantiating experimental investigations summarized in the section on "Experimental Confirmation of Theory," and we would like to re-emphasize this possibility now.

We wish to express our sincere appreciation for Mr. Borden's discussion and references. The induction galvanometer, as discussed by Gilbert, is a somewhat different instrument than the a-c galvanometer considered in this paper. However, in its own right the induction galvanometer is both an interesting instrument and of the same genre as the a-c galvanometer, and our discussion of the latter will, we believe, enable easier grasp of the fundamental principles of operation of the former. Again, the patents cited by Mr. Borden contain some very interesting material. And finally, Mr. Borden's 1923 paper clearly evidences, as Mr. Borden remarks in his discussion, the lack of ready information on the a-c galvanometer, or of references to pertinent sources thereof, in the texts on electrical measurements then available; and, as we point out in our paper, this situation is yet unchanged! Thus we warmly welcome Mr. Borden's strong substantiation of the claim advanced early in our paper—namely, that this paper both fills a gap in the present literature of electrical measurements and advances specific information enabling ready use of the a-c galvanometer.

Thermocouple-Type Ammeters for Use at Very High Frequencies

O. G. McANINCH
MEMBER AIEE

THE measurement of currents at the higher frequencies is complicated by the ever-magnified inductances and capacitances associated with the measuring equipment. On direct currents, permanent-magnet moving coil instruments are used extensively, and here the main consideration is the series resistance of the instrument. Instruments for lower frequency alternating currents generally make use of the electrodynamic or some form of moving iron measuring system. Here again the series resistance of the instrument is important but, in addition, the series inductance of the necessary coils becomes a factor. As the frequency is increased the effects of the inductance also increase, and in the middle audio-frequencies these effects become formidable.

This paper reviews problems encountered in the measurement of current at the higher frequencies, and discusses the performance of ammeters tested in the very-high-frequency range. It also presents the theory, construction, and performance of an experimental wide-frequency-range ammeter having superior characteristics in the very-high-frequency band. It is hoped that the information which is presented will be helpful to designers in fufilling future needs for superior ammeters in the very-high-frequency range.

Instruments for Moderately High Frequencies

In the lower radio frequency range, various thermal conversion devices have been developed for current measurement. In these instruments the current flowing produces heat, and the resulting temperature effect is used as a measure of current magnitude. By this means the coils of wire, with their associated high induct-

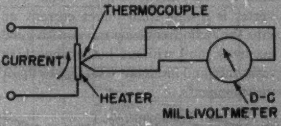


Fig. 1. Elementary thermocouple ammeter diagram

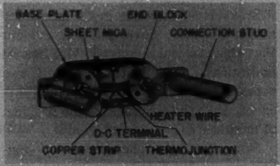


Fig. 2. Conventional thermocouple for radio-frequency instrument

---

Paper 54-170, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Winter General Meeting, New York, N. Y., January 18–22, 1954. Manuscript submitted August 3, 1953; made available for printing December 10, 1953.

O. G. McAninch is with the General Electric Company, West Lynn, Mass.

The author wishes to acknowledge the contributions of his associates in the preparation of this paper, particularly those of T. A. Rich, H. R. Meahl and P. C. Michel of the Company's General Engineering Laboratory, who were primarily responsible for the early development work on low impedance thermocouple ammeters.

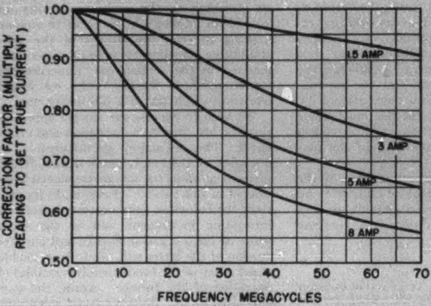


Fig. 3. Correction factor curves for conventional thermocouple ammeters, based on skin-effect calculations

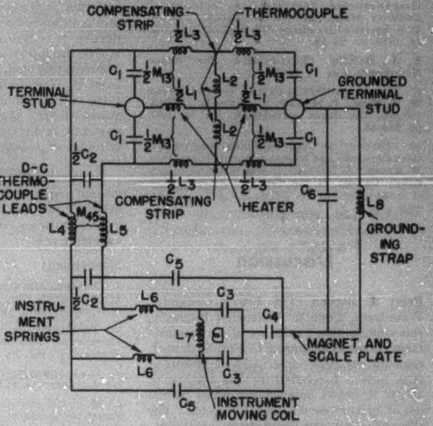


Fig. 4 (right). Internal circuit of conventional thermocouple ammeter at high frequencies

ance, are eliminated from the a-c circuit, and it is possible to produce measurable temperatures in simple heater wires without bothersome high resistances.

The most satisfactory of these thermal converters is of the thermocouple type. Basically (Fig. 1) this consists simply of a short length of wire through which the current to be measured is led. A thermocouple junction is secured to the center of the heater wire and measures the temperature rise produced by the current flow. This unit is the primary detector of the high-frequency current, and operates in conjunction with a sensitive d-c millivoltmeter, which is the end device and indicates the high-frequency current in terms of the thermocouple output. Fig. 2 pictures a typical thermocouple converter of this type, which has been provided with cold junction thermal compensation to maintain a stable reference temperature for the thermocouple.[1] This converter is usable from direct current up into the radio frequency range with good accuracy, and without bothersome high resistance and inductance factors to upset measurements.

Problems Introduced by Necessity of Making Measurements at Higher Frequencies

Still further increases in the frequency range, however, have made it evident that this thermal converter, also has its limitations. Mainly these are of two types. The first is the well known skin effect, while the second is the old problem of increased impedance effects.

Skin effect, the phenomenon of nonuniform current distribution in conductors at high frequencies, causes the heater resistance to increase with increased frequency. This increases the heating for a given current level, and causes the instrument to read high. Hence it becomes necessary to apply a multiplying factor to correct instrument readings, and this factor varies with frequency. Some representative correction factor curves, based on skin effect computations,[2] are shown in Fig. 3. As can be noted from these curves, the correction is greater for instruments of higher current ratings, since the change in resistance occasioned by skin effect is greater in their larger sized heater wires. In certain commercially available instruments the solid heater wire has been replaced by a thin-walled tubular heater which serves to reduce skin effect errors substantially.

The other problems impedance effect, results from inductances and capacitances associated with the converter which are negligible at low radio frequencies. Effectively, at higher frequencies, there is associated with the heater wire both a series inductance and a shunt capacitance. For example, a typical converter having a d-c resistance of the order of 0.03 ohm may have an impedance of the order of 60 or 70 ohms at 200 megacycles (mc). This means that the voltage drop between terminals of such an instrument, an ammeter, is of the order of 300 to 350 volts with 5 amperes flowing. In addition to this disturbing condition there are many stray and shunt capacitances through which such a voltage may force appreciable currents in the frequency range to 200 mc.

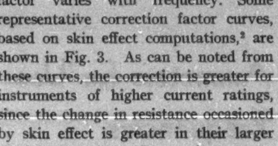


Fig. 5. A low-impedance thermocouple ammeter for use from 0 to 200 mc. View at right illustrates unique concentric terminal arrangement

The Conventional Instrument at Higher Frequencies

An analysis of a typical instrument of the conventional type shows that the actual high-frequency circuit between instrument terminals, neglecting resistance, approximates that shown in Fig. 4. This circuit consists essentially of four elements:

1. The thermocouple unit, consisting of the heater wire L_1, compensating strips L_3 with their shunting capacitances C_1, and the thermocouple wires L_2.
2. The d-c instrument leads, represented by L_4, L_5, and C_2.
3. The d-c instrument, consisting of control springs L_6, moving coil L_7 with its shunt capacitance C_3, and the coil to magnet capacitance C_4.
4. The inductance L_8 of the magnet, scale plate, and grounding strap, and their shunt capacitances to ground C_6.

There are in addition the various mutual inductances as shown.

The approximate nature of the circuit is apparent when it is noted that most of the parameters are actually distributed, rather than lumped. This fact, in addition to the complexity of the lumped parameter circuit and the difficulty of

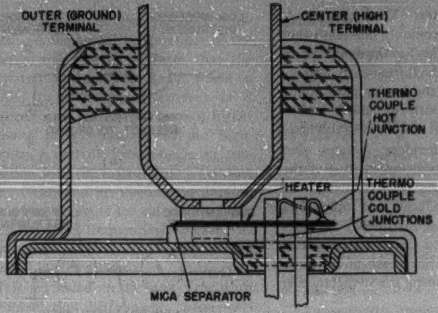


Fig. 6. Cross-section view showing construction of the low-impedance thermal converter

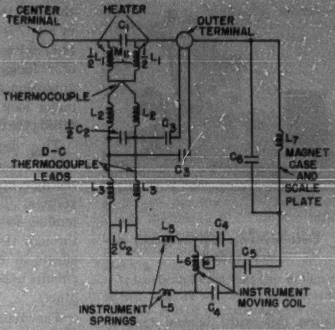


Fig. 7. Internal circuit of the low-impedance thermocouple ammeter at high frequencies

estimating or calculating accurately these parameters, led to the conclusion that a general qualitative approach with experimental verification on actual sample instruments would be preferable to a complete mathematical analysis of the approximate equivalent circuit. Examination of Fig. 4, however, in conjunction with other known information, serves to point out the error-producing effects.

First, the shunt capacitances make possible the diversion of current from its intended path. Second, it is evident that there is a major loop circuit starting at the thermocouple junction, proceeding through the compensating strips, the d-c instrument leads in parallel, via the moving coil to magnet capacitance, to the low terminal stud through the magnet and grounding strap, and back to the starting point through one-half of the heater wire. When it is considered that approximately half of the terminal-to-terminal voltage, which can be of the order of 100 volts or more, is included in this loop, it is evident that heating of the thermocouple wires by the high-frequency current flowing in this circuit will be a considerable factor.

In addition, it is seen that there is the possibility of resonance in all of the many loops (including the major loop described) at various frequencies. Such resonances will produce localized heating effects having the sharply peaked characteristics associated with the resonance phenomena. There is also the possibility of induced circulating currents in the various loops produced by the linkage of high-frequency fields with these loops.

Since the thermocouple measures total heat, and does not distinguish its source, all of these heating effects cause errors in indication. The major sources of errors appearing in the thermocouple instrument at high frequencies may therefore be listed as follows:

1. Resistance rise resulting from skin effect.
2. Extraneous heating of the thermocouple wires as a result of high frequency currents flowing in them.
3. Resonance effects in various of the loop circuits at various frequencies.

An Instrument for Use at Higher Frequencies

Consideration of these sources of error has led through various development stages to an experimental sample of a low-impedance thermocouple type of instrument in which such errors have been reduced greatly. This instrument, illustrated in Fig. 5, has a useful range of 0 to 200 mc, with minimum errors, and without troublesome resonance or impedance effects.

In this instrument a short concentric terminal arrangement replaces the usual side-by-side terminal studs which are usually separated by an inch or more. This configuration reduces the internal circuit length of the instrument, and assists in impedance reduction. It likewise serves to shield the d-c instrument parts from the high-frequency current field, and to prevent the series induction of high-frequency voltages in parts of the d-c instrument circuit. The outer terminal is grounded to the case, as are all other large metal masses in the instrument.

The heating element is made from a thin, flat ribbon, rather than from the usual round wire, to reduce the skin effect error as much as possible. For minimum skin effect, the heater should be very thin; in fact, the thinner the better. Practically, however, it is necessary to

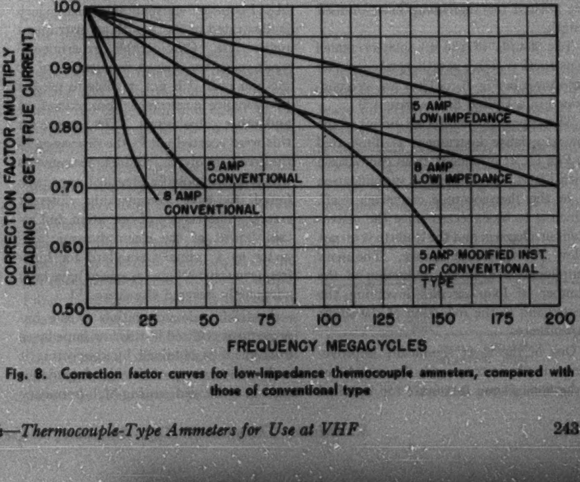


Fig. 8. Correction factor curves for low-impedance thermocouple ammeters, compared with those of conventional type

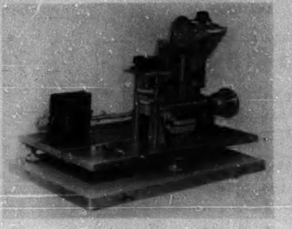


Fig. 9. Electrodynamic high-frequency current standard

compromise this with thermal requirements (which are dependent also on current ratings), so that it is not possible entirely to eliminate the effects of resistance change with frequency. To lower the series impedance of the instrument further, the inductance of the heater is lowered by folding the flat strip heater back on itself, forming a hairpin like structure whose two halves are separated and insulated by a thin sheet of mica.

The path of high frequency current is in the center terminal, through the strip heater, and out the outer terminal. See Fig. 6. The thermocouple junction itself is welded to the mid-point of the heater strip near the fold, and the cold junctions are welded to two small pins which are thermally intimate with the plate connecting the inner and outer terminals, but electrically insulated from it. This arrangement results in d-c thermocouple outputs equivalent in magnitude to those of the conventional arrangement.

To prevent any possible series induction in the d-c instrument loop circuit, the two leads from the thermocouple cold junctions to the instrument armature are twisted together along most of their length.

The details of the completely sealed construction of this new thermal converter are shown in Fig. 6. (They are also contained in a United States patent.[3])

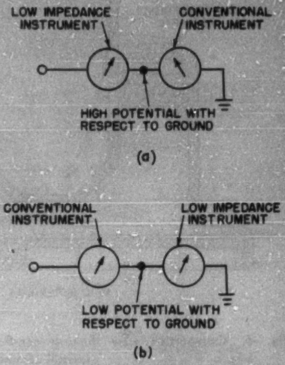
The approximate internal circuit of the low-impedance instrument is illustrated in Fig. 7. It is evident, from comparison with Fig. 4, that the major simplification is in the thermocouple converter itself, which is to be expected since the instrument mechanism and the leads to it do not allow any significant change. The main simplification is in the reduction of the number of loop circuits involving the heater and in the reduction of the shunt capacitances.

One of the most significant improvements in the instrument is the impedance reduction from terminal to terminal through the heater. The value of L_1 in Fig. 7 is approximately one-tenth that of L_1 in Fig. 4. This has the double effect of reducing the high-frequency voltage appearing across the heater part of the major loop previously discussed, and of raising the resonant frequency of this loop.

This design therefore reduces skin effect, reduces extraneous heating of the thermocouple wires by reducing the voltages (both induced and otherwise) tending to send high-frequency currents through them, and decreases resonance effects by eliminating some of the loop circuits and decreasing the inductance in others.

The actual effectiveness of this approach is demonstrated by the fact that instruments of this type are approximately one-tenth the impedance of the conventional type (of the order of 6 ohms), that they are consistently inductive (approximately 0.005 microhenry) with no resonance points from 0 to 200 mc, and that their over-all errors, as indicated in Fig. 8, are lower than those of other thermocouple instruments. From Fig. 8, for example, it is seen that a low-impedance 8-ampere instrument has a correction factor of 0.88 at 50 mc, and 0.70 at 200 mc, while in a 5-ampere instrument the factors are 0.95 at 50 mc and 0.80 at 200 mc. The differences between the 5- and 8-ampere correction factors are due mainly to the necessity of using a thicker heater in the 8-ampere rating, thereby increasing skin-effect errors.

For comparison, Fig. 8 includes correction factor curves on comparable instruments of conventional type, and of conventional type improved to lower skin-effect errors only. The correction factors on each instrument were obtained by comparing it with an electrodynamic high-frequency current standard, which is illustrated in Fig. 9 with an instrument under test. Considerable experimental work indicates that at frequencies of 50 mc and over it is not possible to predict performance accurately from skin-effect and other calculations. The reasons for this were pointed out in the discussion of the errors involved. In fact, instruments which are supposedly identical have been found to have considerably different errors. The theory of operation and establishment of the electrodynamic ammeter as a primary standard for high-frequency current measurement have been adequately covered elsewhere.[4,5,6]



(a)

(b)

Fig. 10. Circuit diagram showing the importance of low impedance in a thermocouple-type instrument and the significance of ammeter location in measuring circuits

A striking demonstration of the improvement realized in this low-impedance instrument is obtained by connecting it and an instrument of the high-impedance type in series and passing high-frequency currents through them. When connected as shown in Fig. 10(A), with the low impedance instrument at a higher potential, that instrument will read (with correction factor applied) considerably higher than the high-impedance ammeter. When the positions of the instruments are reversed, as in Fig. 10(B), and the proper correction factors applied to both instruments, their readings will be found to agree. The reason for the difference is the diversion of currents from high potential parts of the circuit to ground, thus allowing current to flow through one instrument and not the other.

Such a demonstration as this also brings out the point that the current indicated by the low-impedance instrument is that which flows into the center terminal, and not a fictitious current indication produced by the variety of error-producing factors previously outlined.

Conclusion

By the proper application of these principles in the design of a low impedance thermocouple-type ammeter, it is possible to make accurate current measurements at frequencies from 0 to 200 mc with a minimum of disturbance of circuits, as well as to determine the relative magnitudes of currents. Since the variation of multiplying factors at different frequencies is minimized by these design principles, it has been proved possible to calibrate development sample instruments of low impedance design so that they may be used accurately over a predetermined and extensive frequency range without correction.

Consideration of the theoretical aspects and the actual performance of the experimental sample indicates that the suggested low-impedance design would be limited to a maximum current of 8 amperes, and that for measuring higher currents it is necessary to use high-frequency current transformers in conjunction with the instrument. This and other techniques involved in the application of instruments for current measurement at these frequencies represent an important subject in themselves not covered in this article.

It is concluded, on the basis of the study described, that the frequency range of the thermocouple-type ammeter can be substantially extended by adoption of suitable design principles.

References

1. The Compensated Thermocouple Ammeter, W. N. Goodwin, Jr. *AIEE Transactions* (*Electrical Engineering*), vol. 55, Jun. 1936, pp. 23–33.

2. *Bulletin No. 74*, United States Bureau of Standards, Washington, D. C.

3. Rich, Meahl, Michel. United States Patent No. *2,199,247*.

4. A High Frequency Electrodynamic Ammeter, P. C. Michel. *Thesis*, Yale University, New Haven, Conn., 1935.

5. An Electrodynamic Ammeter, H. M. Turner, P. C. Michel. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 25, pp. 1367–74, Nov. 1937.

6. A Bearing Type High-Frequency Electrodynamic Ammeter, H. R. Meahl. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 26, June 1938, pp. 734–44.

7. Thermocouple Ammeters for Ultra-High Frequencies, John H. Miller. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 24, pp. 1567–72.

8. Frequency Errors in Radio-Frequency Ammeters, J. D. Wallace. A. H. Moore. *Proceedings*, Institute of Radio Engineers, New York N. Y., vol. 25, pp. 327–39, March 1937.

◆

Discussion

H. R. Meahl (General Electric Company, Schenectady, N. Y.): It may be of interest to know that arbitrarily setting a lower limit of 0.8 on the frequency correction factor allows the use of a 1.5-ampere instrument over the frequency range of 0 to 300 mc per second, and a 3-ampere instrument up to 240 mc.

It has been found that this design resulted in instruments which can be operated in extremely strong electromagnetic fields, such as those existing within a foot of conductors carrying 100 amperes at 810 kc without sacrificing accuracy.

Pulse Relaxation Amplifier—A Low-Level D-C Magnetic Amplifier

R. E. MORGAN
MEMBER AIEE

J. B. McFERRAN
ASSOCIATE MEMBER AIEE

DEMAND for stable amplifiers to work at extremely low signal levels has attracted the attention of all who design and build amplifiers. The pulse relaxation amplifier (pra) is an outgrowth of this demand. It is essentially a magnetic amplifier requiring no rectifiers and operating from a pulse power supply. Preliminary tests indicate a zero drift level of less than 10^{-16} watt over a temperature range from −70 to 140 degrees centigrade. This is compared to 10^{-8} watt attainable with conventional amplistats (self-saturating magnetic amplifiers). They are temperature-limited primarily because of their rectifiers. The two greatest sources of drift in amplistats, variation in rectifier and magnetic-core characteristics, have been eliminated in the pra.

The pra is not an amplistat, since by definition the latter achieves self-saturation by means of rectifiers in the output circuit. The pra employs the technique of short pulses separated by long relaxation periods. Flux reset in the core occurs during the relaxation periods. The pra operates from a pulse power supply and the pulse supply, in turn, from a sinusoidal power source. The magnetic core of the amplifier is driven into saturation by the power pulses. During the relaxation period the core magnetization settles back to a level dictated by the control signal. The output power is determined by the amount of pulse remaining after saturation. Push-pull operation is achieved using a single magnetic core by comparing the alternate positive and negative pulses. Approximately the same gain and speed of response are realized in the pra as in the amplistat.

Drift problems in the conventional amplistat are largely attributed to forward rectifier drop, reverse rectifier leakage, and variations in the magnetic characteristics between cores. By eliminating the rectifiers, reducing the number of operational cores to one per amplifying stage, and by operating this core in push pull, the pra has considerably less opportunity for drift. Regulations in the sinusoidal power supply, both in magnitude and in frequency, are compensated. The circuitry is so devised that variations in saturation flux density and hysteresis loop width have negligible effect. Changes in hysteresis loop slope affect the gain only, not the level of drift.

The pra is basically a low-level device. As such, several stages of pulse relaxation amplification can be used to amplify a low-level signal, bringing it to the microwatt level. Here conventional amplistats can be used to boost the power to higher output levels without further effect upon the drift. The concept of the pra suggests a wide variety of new functions which may be performed by magnetic amplifiers. These cover the fields of optical pyrometry, precision calorimetry, spectroscopy, geophysics, meteorology, differential thermometry, etc. Wherever extreme sensitivity is desired, the pra has an application.

Theory of Operation

Operation of PRA Input Stage

A simplified circuit of the pra input stage is shown in Fig. 1. Unlike a half-wave amplistat, this circuit employs no rectifiers. Also, the core size is much smaller than is normally used in amplistats. Fig. 2 represents an idealized hysteresis loop (four straight lines) of the amplifier core characteristic. The axes

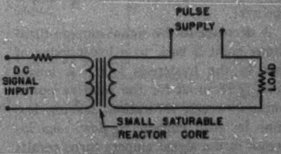


Fig. 1. Simplified pra input stage

Paper 54-198, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE North Eastern District Meeting, Schenectady, N. Y., May 5–7, 1954. Manuscript submitted February 8, 1954; made available for printing March 18, 1954.

R. E. Morgan and J. B. McFerran are with the General Electric Company, Schenectady, N. Y.

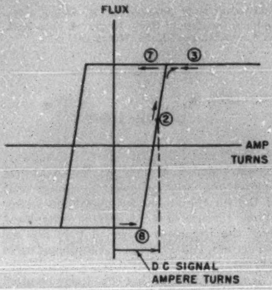


Fig. 2 (left). Hysteresis loop of pra input stage reactor

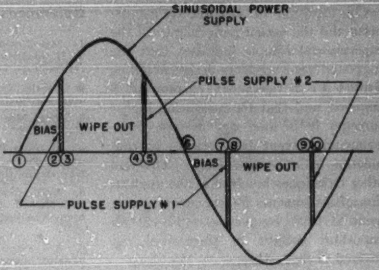


Fig. 3 (right). Pulse and bias supplies relative to sinusoidal power source

of the loop are expressed in flux versus ampere-turns, since no variation in core dimensions is introduced in this report. The encircled numbers of the loop establish a "time" correlation between the magnetic state of the core and the sinusoidal alternating voltage of the main power source. This is illustrated in Fig. 3 which shows the pulse supply relative to the main sinusoidal power source. Each number represents a point in time. Pulse width is 2 degrees of the sinusoidal supply period.

Following the operation in steps, start with point 2 on the right-hand side of the hysteresis loop illustrated in Fig. 2. This starting point is established by the d-c signal ampere-turns. The reactor is designed so that the pulse supply volt-time integral will swing the core flux over the entire range of the hysteresis loop. Therefore, the first portion of the pulse time is used to raise the flux level from point 2 to saturation. Full load current will then flow for the remainder of the pulse. The core remains saturated at point 3 until the beginning of the negative pulse. The total duration of the negative pulse is required to drive the core from positive to negative saturation. As negative saturation is just reached by the end of the pulse, no negative output current is obtained. The d-c signal returns the flux level in the core to point 2. However, a definite period of time is required for the flux to accomplish this shift in level. This is known as the relaxation period and represents the time between the end of one pulse and the beginning of the next.

A decrease in d-c signal ampere-turns drops point 2 back to a lower value of flux in the hysteresis loop. A longer time is required for the positive power pulse to reach saturation, leaving less time for load current to flow. An increase in d-c signal ampere-turns results in less time for the pulse to reach saturation and hence more time for load current to flow. In both instances the negative power pulse is completely absorbed in swinging the core from positive to negative saturation. A negative d-c signal causes the roles of positive- and negative-pulsed outputs to be reversed. A plot of output current, averaged during the pulse periods only, versus signal current provides the characteristic shown in Fig. 4.

BIAS

Fig. 4 indicates that the pra gain characteristic is bidirectional about the origin of the input and output axes. A low gain section in the characteristic appears at the origin because of the hysteresis loop width. This may be corrected by the addition of a cross-bias winding and supply to the pra. The bias must operate alternately positive and negative. It is applied most conveniently from a sinusoidal a-c supply. The bias supply adds ampere-turns to the core in addition to the signal winding ampere-turns. Fig. 5 illustrates the use of a-c bias. It is adjusted so that at points 2 and 7 the flux is set halfway between positive and negative saturation. This results in the flow of load current during both positive and negative pulses with a net output of zero. When a small d-c signal is applied in addition to the bias, points 2 and 7 are shifted accordingly, and a net output occurs as a result of oppositely directed control on both positive and negative pulses. Fig. 6 shows the gain characteristic with bias applied.

It is important to note that the net output represents the difference between the average currents during the pulses only. The small currents which occur during the relaxation period are not passed on to the following amplification stages. These currents are the result of the flux resetting voltages induced in the power winding. They may be ignored. A circuit composed of linear resistive and nonlinear reactive components, and supplied from an a-c supply having a zero average component, cannot produce rectification. The net current averaged over

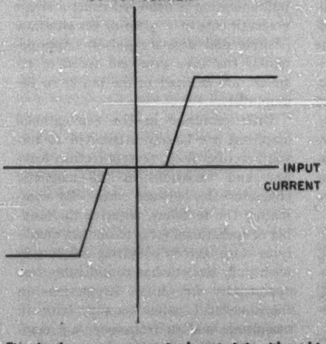


Fig. 4. Input stage pra gain characteristic without bias

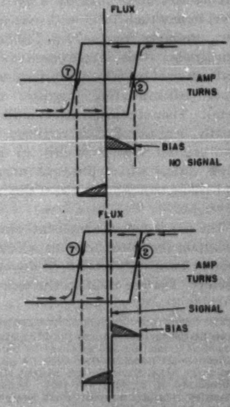


Fig. 5. Pra input stage hysteresis loop with a-c bias

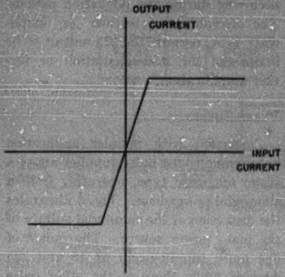


Fig. 6. Pra input stage gain characteristic with bias

the pulses and relaxation periods must be zero. Although a net current difference may occur between the pulses, this is compensated by an equal but opposite difference during the relaxation periods. Since the pulse magnitude is many times larger than that during the relaxation period, only the pulses are passed to the following stages.

So far, loading effects on the pra circuit through transformer coupling have been ignored. In a practical application of the pra principle, a means must be provided for achieving operation without loading through transformer coupling. The circuit for accomplishing this is shown in Fig. 7. Two cores are used instead of the previous single core. The added core acts in an auxiliary capacity, not in a manner whereby its characteristics are compared with those of the main core. The power-winding circuit includes a filter element composed of resistor and capacitor. This acts as a high impedance to the slowly varying signal and bias supply induced voltages. However, to the pulse supply voltage, the filter is effectively a short circuit. The signal windings and bias windings are connected in such a manner that the induced voltages caused by one supply cancel out in the windings of the other.

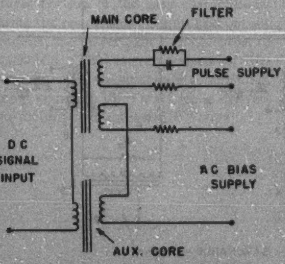


Fig. 7. Input stage pra circuit

Operation of PRA Interstages

For additional stages of pulse relaxation amplification, an additional pulse power supply operating in time displacement relative to the original pulse supply is needed. This permits one supply operating in conjunction with the power winding of the previous stage of pulse relaxation amplification to set the signal flux level. A short time later the second pulse supply acting through the power winding causes the core to saturate and power to be delivered to the following stage signal winding. The power winding of this stage is connected to the original pulse supply. Summarizing, the interconnection of several stages of pulse relaxation amplification requires two pulse power supplies which are connected between alternate pra stages and furnish their pulses in alternate sequence. Fig. 3 shows the pulses of both supplies in time relation to the sinusoidal power. The pulses of one supply are between points 2 and 3 and between points 7 and 8. Those of the other supply are between points 4 and 5 and between points 9 and 10.

Consider an interstage of pulse relaxation amplification whose control signal is, instead of the steady direct current of the input stage, a pulsed signal. Other than this the circuit of Fig. 1 applies and may be used for a description of the operation. The magnetic changes of the core are shown on the hysteresis loop in Fig. 8. At point 2 the flux is at negative saturation and zero ampere-turns. This is the residual condition left by the preceding negative power pulse. The signal pulse applied at this point has sufficient amplitude and duration to make the transition to point 3. At point 3 the signal pulse is removed. The flux remains unchanged until the beginning of the power pulse. At point 4 the power pulse is applied, driving the core into positive saturation. Only a portion of the power pulse is required for this purpose, leaving the remainder for output. After the power pulse is over, the core drops back to the residual condition of zero ampere-turns at positive saturation. A similar sequence is followed by the negative signal and power pulses.

The loading effects caused by transformer coupling of the signal and power windings must also be considered in the pra interstages. Fig. 9 shows a pulse relaxation interstage amplifier designed to partly eliminate the loading effect of the pra and partly take advantage of it. Two cores, a main one and an auxiliary, are used. Both cores have common signal and bias windings. The power winding links both cores, while a "wipe-out" winding links only the auxiliary core. Induced components of signal voltage appearing in the power windings of the main and auxiliary cores buck out one another. The bias and wipe-out voltages are applied to their respective windings through large-value resistances, so that these circuits contribute very little to the loading of the other windings and their circuits. Although these circuits are loaded by the presence of the signal-winding circuit, their voltages are applied over long periods. The necessary flux reset is realized in this manner.

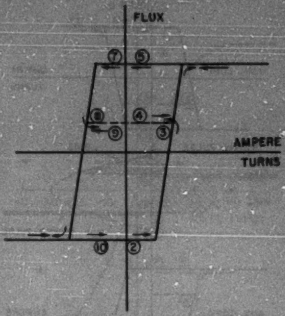


Fig. 8. Pra interstage hysteresis loop

Referring to Fig. 10 a hysteresis loop of the auxiliary core is drawn with the ampere-turns scale reversed so that a simultaneous comparison between it and the main core loop can be made. The bias drives the auxiliary core further into negative saturation between points 1 and 2, while the main core is driven into positive saturation. As the signal is applied,

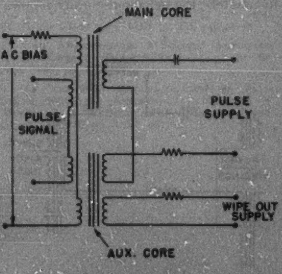


Fig. 9. Pra interstage circuit

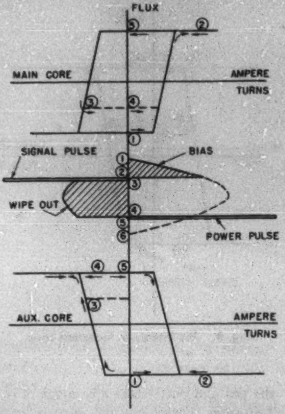


Fig. 10. Time flux condition for pra interstage

the same flux change occurs on both cores, but in opposite directions. After the signal pulse there is no flux change in the main core, between points 3 and 4, but the auxiliary core is driven into positive saturation by the wipe-out voltage. When the positive power pulse occurs, it finds the auxiliary core already in saturation. The auxiliary core therefore presents no impedance to the load circuit during the power pulse, but only during the signal pulse.

The capacitor in the power winding becomes partly charged during a power pulse. After the pulse has ended the capacitor retains its charge. In attempting to discharge, the flow of current is reversed from that occurring during the pulse. But this reversed current is opposed by the impedance of the power winding. The capacitor voltage decays only slightly. Meantime, sufficient volt-seconds are applied to the core by the capacitor to cause it to swing from its saturated state to saturation of the opposite polarity. Unfortunately, the signal pulse which follows finds the core already in saturation. It can do no more than drive the core further into saturation. Thus all control is lost. By changing the polarity of the windings, control is regained. However, the sense is now altered. The signal pulse drives the core in the opposite direction and from the opposite state of saturation than was previously done. The power pulse, however, is still acting in the same direction. Conceivably, a large signal pulse could return the core to the state of saturation existing before the capacitor volt-seconds were applied. But this core state, in the previous signal connection, was obtained from a small signal pulse. Conversely, a small signal pulse in the new mode of operation sets the core to the same point as a large signal pulse in the previous mode. Thus a small signal results in a large output and a large signal results in a small output. The sense has been changed. Since the pra operates in push pull, the net output between positive and negative pulses is still the same as that obtained with the previous connection. Added stability is realized by this new mode of operation.

Additional Circuit Considerations

Multistage PRA Operation

Pra stages may be cascaded together until a signal is increased to a sufficient level that it can be amplified by more conventional magnetic means. Fig. 11 illustrates the interconnection of two interstages of pulse relaxation amplification. Transition from the pra to an amplistat requires a discriminator between the two units. Although the pra positive and negative pulses may differ in width as a function of the control signal, the average current over both the pulse and relaxation periods must still be zero. Thus the discriminator must take the difference in currents occurring during the pulse periods, and must eliminate those occurring during the relaxation periods. The discriminator, rectifier or amplistat[1] type, permits the pra output to be intepreted for instrumentation or further amplification.

Pulse Supply

Consistent with popular application requirements, the pulse supplies are of a static magnetic type, operating from a sinusoidal power base. Fig. 3 illustrates the dissection of the sinusoidal voltage of the main power source. The pulses of one pulse supply occur at 45 and 225 degrees, while those of the other pulse supply occur at 135 and 315 degrees. The relaxation period between pulses is not critical in its duration. Each pulse is approximately of 2 degrees' duration (100 microseconds in a 60-cycle source). Much of the remaining portion of the sine wave is used to obtain bias and wipe-out voltages.

Fig. 12 shows the circuit used to obtain the pulse voltages. It is composed principally of transformers which are alternately saturated positive and negative. Returning to the sine wave of Fig. 3, the first 45 degrees of the sine wave voltage are held off by X_3. When X_3 saturates, the power is transferred from X_3 to the circuit made up of R_1, R_2, and X_4. X_4 is a transformer having a small core and few turns. It saturates after only 2 degrees of angle, leaving the remaining portion of the half-cycle to be dissipated in resistor R_1. The secondary voltage of X_4 constitutes one pulse supply. The other supply pulse is achieved by similar means from X_9. X_8 is designed to hold off the first 135 degrees of the sine wave voltage, after which the circuit composed of R_6, R_7, and X_9 receives the power. X_9 saturates rapidly so that its secondary voltage constitutes the other pulse supply. Also shown in Fig. 12 are the bias voltages and their wave forms.

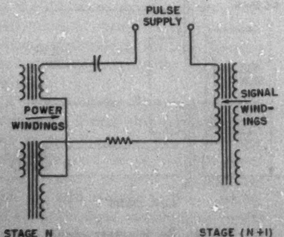


Fig. 11 (left). Interconnection of pra interstages

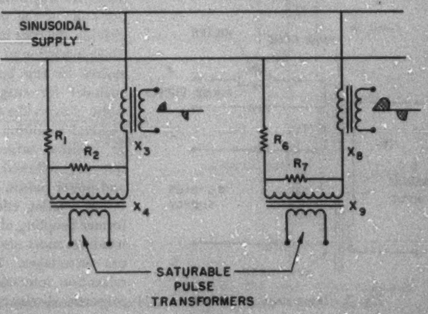


Fig. 12 (right). Power supplies

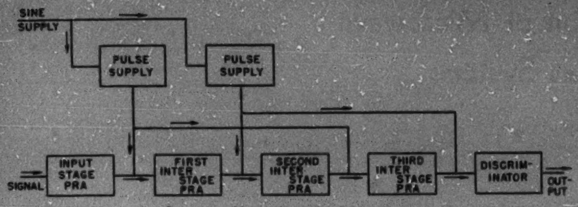


Fig. 13. Block diagram of 4-stage pra

The outstanding advantage of a magnetic-type pulse supply for the pra lies in the fixed volt-second integral provided by the pulse supply transformers. Regulation of voltage and frequency in the sinusoidal supply may alter the pulse shape, but not the pulse integral. Since the pra cores each require fixed volt-second values to swing the flux over their saturation range, the fixed pulse integral contributes much to the amplifier stability.

Characteristics of 4-Stage PRA

GAIN

A multistage pra with a d-c input and a d-c output was built and tested. It consisted of an input stage, 3 interstages of pulse relaxation amplification, pulse power supplies, and a rectifier discriminator for output. Fig. 13 shows it in block diagram. It was designed to operate from a 115-volt 60-cycle power source. The resistance of the input stage was 50 ohms while that of the discriminator output was 200,000 ohms. Deltamax toroidal cores of half-inch diameter by 0.0005-square-inch cross section were used in the pra stages. Fig. 14 shows an over-all gain characteristic for the amplifier in terms of average volts output versus average volts input. Inasmuch as the input and output resistance levels are different, the over-all amplifier voltage gain has less significance than an over-all power gain. If such is computed from the linear portion of the gain characteristic using the resistances of input and output circuits, the power gain is of the order of 2,000,000. The linear portion of the amplifier will accommodate a maximum input signal of 10^{-11} watt while delivering an output in excess of 10^{-6} watt.

SPEED OF RESPONSE

Speed of response in the pra is a function of the input stage time constant and the interstage pulse delays. The time constant of the signal winding in the input stage is inherently low. This is principally the result of using small core sizes, which usually give quicker response. Between stages of pulse relaxation amplification, quarter-cycle pulse delays are involved. That is, the second stage will not respond before its pulse, which occurs 90 degrees after the first stage pulse. Pulse delay is also involved in the input stage of the pra. However, here it may take any value from 0 to 360 degrees of the sinusoidal supply frequency. Assuming an average input stage delay of 180 degrees, the response of a multistage pra circuit is the sum of the time constant of the input winding and the product of the number of stages plus one times a quarter period of the sinusoidal power supply.

DRIFT

The results of preliminary drift tests are presented here. These tests were made to estimate the influence of temperature variations and power supply variations on output and gain of the pra. The drift level was determined by observing the output of the amplifier with the input signal held at zero. An input signal, which when amplified gave the same magnitude of output as that given by the drift level, defined the minimum signal level of the amplifier. Recordings were made over a 2-week period of the pra output. Power supply voltage was varied ±30 per cent and temperature for the first stage was cycled between -70 and 140 degrees centigrade. Maximum drift during this period was 10^{-16} watt.

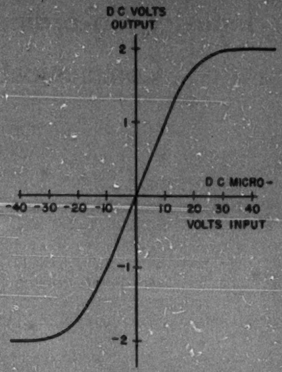


Fig. 14. Gain characteristic of 4-stage pra

The gain characteristic showed complete continuity. A check of gain stability showed variations of 10 per cent for power supply frequency variations of ±10 per cent.

Conclusion

Circuit constants and variations upon the circuit types have for the most part been omitted. The purpose of this report has been to present a novel technique of magnetic amplification, therefore detailed circuit information has been minimized. The pra realizes the functions of signal reset and power gain, requirements of any magnetic amplification scheme. Its significant application lies in a region of signal sensitivity not heretofore available in magnetic amplifiers

References

1. THE EXTENSION OF AMPLISTAT PERFORMANCE BY A-C COMPONENTS, R. E. Morgan, H. M. Ogle, V. J. Wattenberger. *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 986–91.

2. THE AMPLISTAT AND ITS APPLICATION, H. M. Ogle. *General Electric Review*, Schenectady, N. Y., vol. 53, Feb. 1950, pt. I, pp. 32–35, Aug. 1950, pt. II, pp. 41–44, Oct. 1950, pt. III, pp. 41–46.

3. THE AMPLISTAT—A MAGNETIC AMPLIFIER, R. E. Morgan. *Electrical Engineering*, vol. 68 Aug. 1949, pp. 663–67.

4. AN AMPLISTAT REGULATOR APPLIED TO A 5-KVA 400-CYCLE AIRCRAFT INVERTER, R. E. Morgan, J. A. Walley. *AIEE Transactions*, vol 69, pt. II, 1950, pp. 1,243–48.

No Discussion

Industry Co-ordination of Microwave Communications Systems

VICTOR J. NEXON
ASSOCIATE MEMBER AIEE

THIS paper is a discussion of the problems concerned with co-ordinating microwave systems in the various safety and special services to obtain maximum utilization of the microwave spectrum consistent with the economical and technical factors of system usage by present and potential users.

In view of the directive radiation characteristics of microwave antenna systems, co-ordination solutions on a beamwidth basis together with frequency and bandwidth should be found to provide interference-free operation for a maximum number of microwave systems in the various areas.

Need for Industry Co-ordination

There are areas developing in this country where special co-ordination consideration must now be given to future installation of microwave systems to avoid interference. Some of these areas are in the Valley Forge, Pa., Houston, Texas, and Kansas City, Mo., vicinities.

Because of the importance of hilltop locations, these and other areas are rapidly reaching a point of band saturation, especially where such hilltops are few in number and where their positions make them especially suitable. Furthermore, fixed points for communication requirements are often in the same city or in close proximity for different companies in the safety, transportation, and industrial services. Consequently, the beamwidth patterns may overlap and cause need for separate frequencies in the same or other commercial bands.

In the Valley Forge area there are three separate microwave systems, namely for the Keystone Pipe Line Company, the Texas Eastern Transmission Corporation and the Transcontinental Pipe Line Company, all in the 1,850- to 1,990-megacycle (mc) commercial band. In an area of approximately 10 square miles there are five stations—two repeaters with three directions each, one repeater with two directions, and two terminals—requiring a total of 10 2-way radio-frequency (r-f) channels in the 1,850- to 1,990-mc band. To satisfy this requirement all but four frequencies available in this band were used, and it is doubtful that these four can be used in the future because of possible receiver-image interference.

Careful study was made of beamwidth pattern interference in this area and maximum guard angles were used in choosing the frequencies to allow for sidelobe and overshoot interferences from and to antennas having 1/2-power beamwidths of 5 to 6 degrees. Consideration of safety factors for misfocusing of the dipole elements and for beamwidth azimuth movement caused by wind and ice forces on the paraboloids and towers was another reason for choosing separated frequencies to reduce interference possibilities.

In the Houston-Galveston area there are five separate commercial microwave systems, namely for the Trunkline Gas Company and the Transcontinental Pipe Line Company in the 1,850- to 1,990-mc band and for the Humble Pipeline Company, the Texas-Illinois Pipe Line Company, and the Santa Fe Railroad in the 6,575- to 6,875-mc band.

Had these systems been chosen to operate in the same frequency band, careful juggling of the different frequencies would have been necessary to avoid interference. Restricted beamwidth patterns with guard angles may have been necessary to allow use of the same transmit or receive frequencies. Since Houston is a focal station point, it is conceivable that in the near future new systems installed in this area may require the use of one frequency band in the Houston vicinity and another frequency band in outside, less congested areas. At that time special co-ordination survey studies will become a necessity.

In the Kansas City area there are three separate commercial microwave systems, namely for the Platte Pipe Line Company the Sinclair Pipe Line Company, and the Panhandle Eastern Pipeline Company, all in the 6,575- to 6,875-mc band. Although this band is 300 mc wide as compared to 140 mc for the 1,850- to 1,990-mc band, the use of passive reflectors requires four separate frequencies for transmitting and receiving at a repeater station. Requiring at least a 40-mc separation between transmit frequencies and a 120-mc separation between transmit and receive frequencies, systems of this type use more band frequencies, especially if the beamwidth patterns overlap and require separate frequencies.

Kansas City is a typical growing industrial center. Pipe-line companies, together with power companies, railroads, turnpikes, and other eligible industries, are planning microwave systems in or around areas of this nature. Without co-ordination of microwave frequencies and beamwidth patterns, resulting interferences and confusion will greatly reduce the utilization of the bands now available on a shared basis. Other areas which are also rapidly approaching a point of concern for co-ordination in the 1,900- and 6,700-mc bands are around Los Angeles, St. Louis, and New York.

Potential System Application

The possibilities of microwave communication system application are unlimited, and broad in scope. They exist because microwave serves as a main highway for many varied forms of communication required by many different industry and safety groups which are now obliged to share the frequencies available in the commercial bands of 952 to 960 mc, 1,850 to 1,990 mc, 2,110 to 2,200 mc, 2,450 to 2,700 mc, 6,575 to 6,875 mc, and 12,200 to 12,700 mc.

With progress, the number of interested users will grow and so will the number of systems installed, whether the latter are privately owned, leased from the suppliers, or leased as private systems from the telephone utility companies. Proper methods of co-ordination and control must therefore be set up now to allow for the systematic and best use of the limited frequencies available for these broad-band types of microwave applications.

Industry Sharing of Frequencies

At present all licenses granted are on a developmental basis and all frequencies in the commercial microwave bands are on a shared basis. Service eligibility is usually limited to the groups now eligible for regular fixed and mobile radio service. Applications from other groups such as special industrial and service organizations are considered separately on their relative importance and need. Equiva-

---

Paper 54-188, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Northeastern District Meeting, Schenectady, N. Y., May 5–7, 1954. Manuscript submitted January 21, 1954; made available for printing March 8, 1954.

VICTOR J. NEXON is with Microwave Services, Inc., New York, N. Y.

lent service through use of common-carrier facilities by microwave, cable, or wire is a factor in the consideration of applicants not specifically eligible under the Federal Communications Commission (FCC) rules and regulations.

Because the frequency bands available allow for few 2-way r-f channels with adequate bandwidth and because the beamwidth patterns are highly directive instead of omnidirectional or wide in nature, it is unlikely that microwave frequencies will be allocated in the future on a separate industry group basis as in mobile radio.

Licensing Problems

Since licenses are granted on a developmental basis, full standards have yet to be established for microwave systems, and the industry has had to get along on the various technical standards recommended by the manufacturers and laboratories supplying the equipment. Frequency recommendations, frequency stability, power output, bandwidths, and type emission are given for the manufacturer's type of equipment and usually are accepted by the FCC in its developmental grant for the construction and operation of microwave stations.

Microwave communication systems installed to date and those now available from the various manufacturers differ widely as to the type of antenna systems, r-f equipment, frequency stability, r-f bandwidth, and multiplexing. Installations are growing on a "first come, first served" basis in the given area, and later arrivals will find it difficult to fit into the existing unco-ordinated situations involving different frequency separations, different bandwidths, and variations in antenna directivity and sidelobe tolerances.

Because of this growing confusion there is considerable activity on the part of various industry groups in the direction of standardization and establishment of a permanent set of rules and regulations by the FCC. However, before this can be done, a number of factors must be considered, including the effect on existing installed systems, eligibility of other possible users, extra engineering design costs to manufacturers, methods of controlling and recording beamwidths and guard angles, possible bandwidth packages for voice channel requirements, joint industry co-ordination and administration procedures, system licensing instead of station licensing, field survey analyses, and independent certification. These and other factors are under discussion during the "shifting of gears" from temporary to permanent licensing.

Spectrum Utilization

The possibility of frequency assignments based on geometry rather than on the frequency or service category has been discussed by Edwin L. White.[1] By this, he meant the sharing of frequencies with assignment of beamwidth patterns in the same area, thus increasing the number of potential r-f channels in any one area. He indicated that such an arrangement could be successful only if the best engineering techniques were applied to the directive antenna systems, their supporting structures, and also the associated r-f and multiplex electronic equipment.

Let us now consider what some of the system conditions are for which these "best engineering techniques" must strive to realize optimum spectrum utilization.

R-F Bandwidth and Ultimate Multiplex Channels

A desired situation on bandwidth would be one where the assigned bandwidth is the minimum for the ultimate number of multiplex telephone channels authorized for a potential user. This would allow for systematic grouping of the r-f carrier frequencies in each commercial band consistent with the user's present and future needs.

For example, it would make frequency assignments much easier if r-f bandwidth could grow with channel packages of 4, 8, 16, 24, and 30 telephone channels. Channel bandwidths could be set aside in each band for these packages. Potential users would be required to justify their need for telephone capacity at the time of license application.

Maximum Frequency Stability

The highest order of frequency stability, consistent with available tubes and circuitry technique, is desirable. This minimizes the total r-f transmitter deviation from the assigned carrier frequency and allows for narrower receiver bandwidths. Direct crystal control, automatic frequency-controlled circuits, or temperature-controlled oscillators realizing results of at least 0.01-per-cent stability or better should be strived for.

Minimum and Restricted Beamwidths

The use of antenna and reflecting surfaces whose radiated energy is confined within a minimum specified radiated cone is desired to allow azimuth assignments of transmission and reception.

Consistent with practical costs of tower-supporting structures, the maximum 1/2-power beamwidths allowed should be in the order of 4 degrees in the 1,900-mc band and 2 degrees in the 6,700-mc band. Guard angles on each side of the 3 decibel points should be not more than 15 degrees, at which outside points the signal level of sidelobe radiations should be at least 40 decibels down from the main radiation.

Where antennas are mounted on supporting structures which are subject to twist and deflection, a method of measurement giving a continuous record of the azimuth movement of the radiated beam should be provided. All supporting structures should be so designed as not to twist more than the 1/2-power beamwidth under maximum wind conditions for the area. This would be necessary in order to prevent interference between adjacent beamwidth patterns during high wind conditions.

Minimum System Frequencies

Antenna systems should be so designed as to require the use of a minimum of frequencies for transmit and receive at a repeater station. This may be difficult with some antenna systems but should be a requirement for which to strive in all frequency bands.

Practical Co-ordination Problems

Microwave co-ordination in the commercial bands would be much easier of course, if the foregoing desired conditions existed today, but they do not. Through the processes of free competition, pressure from industry groups, self-control, and government guidance, it is believed that a co-ordinated group of microwave systems and area plans will evolve, which will eventually provide for the most efficient use of the spectrum.

Co-ordination today, however, is faced with utilizing the present variety of practices and analyzing applications of present microwave equipment. To co-ordinate system usage on the basis of exclusive frequency assignment in an area would be using the line of least resistance and would not be in the public interest. On the other hand, frequency assignment plans, based on a geographical basis, bring up many practical problems. Let us analyze a few of these practical problems that experiences have brought to light in the past several years, as noted in the following.

Frequency Assignments and Overshoot Interference

Overshoot interference as a result of abnormal refraction phenomena usually makes it necessary to assign alternate frequencies to a line of repeater stations in the same system; when new stations are located in an area by other users, possibilities of overshoot interference from other areas may result if the antenna orientation is approximately in line and the frequencies are the same.

This is contrary to the expected "line-of-sight only" properties of microwave propagation. At times these overshoot signals may be stronger than the direct signal. Polarization and beamwidth changes in the antennas are of some help but do not give complete rejection. In most cases it will be important that initial field surveys be made to determine which group of frequency assignments, which frequency band, and possibly which type of multiplexing would best fit the locations contemplated. Often these locations may be fixed as a result of communications requirements.

Antenna Sizes, Beamwidths, and Path Lengths

Most systems today are designed on the basis of having adequate reserve gain to allow for fading and maintenance margins. This usually means using higher gain antennas for long path lengths. Beamwidths become smaller with increase in antenna gain and larger with decrease. Paraboloid antenna areas increase with gain and naturally the assemblies become more costly with size. For short path lengths, then, small antennas with wide beamwidths up to 20 degrees, having sufficient gain, have been used for both economical and technical reasons.

As the azimuth space in an area becomes occupied it may be necessary that all antenna 1/2-power beamwidths be restricted to the smallest practical width for the particular frequency band. For example, at 1,900 mc a 2-foot paraboloid has a beamwidth of approximately 18 degrees whereas a 10-foot paraboloid has a beamwidth of approximately 3.5 degrees. Similarly, at 6,700 mc the beamwidths are approximately 5 degrees for a 2-foot paraboloid and less than 2 degrees for a 6-foot paraboloid. Co-ordination efforts produce the problem of what to do with antennas already installed with wide beams in congested areas.

Antenna Beamwidth Control and Tower Costs

It may be easy to say that beamwidths should be as narrow as possible to allow for more sharing of frequencies in an area, but will it be practical to keep these beamwidths contained within narrow limits, especially in areas which have more than average wind velocities?

Experience has shown that twist and deflection tolerances of ±1 degree and ±0.5 degree, respectively, are reasonable under conditions of 30 pounds per square foot of wind loading and 1/2 inch of ice loading, provided the tower heights are not much more than 300 feet.

Measuring and logging the azimuth movement of an antenna and its supporting structure to control and allow close adjacent space operation of another beam, however, is a problem. Various methods have been suggested, such as the light beam method which uses a collimated beam of light from the tower to a calibrated graph which is continually photographed for a permanent record of the tower twist and deflection.

Co-ordination will require that an acceptable method of measurement be used or that independent engineering certification of tower rigidity and deflection tolerances be made as a part of beamwidth space assignments.

Back-to-Back and Reflection Interferences

The radiating and rejection characteristics of paraboloid antennas and passive reflectors are such that sidelobe interferences are possible. Unexpected received signals from reflecting hills or buildings or even from similar strong radiating antennas, 180 degrees away, can often upset what appears to be a logical frequency assignment for the receiving station. Changes in polarization and beamwidth offer some help in reducing this type of interference.

If possible, it would be desirable that antenna reflecting surfaces be improved and that the assemblies be more accurately tuned or focused for each particular frequency. Reduction of sidelobes would assist in lessening the guard angles between adjacent beams of the same frequencies.

In the 1,900-mc band, for example, at a backbone and sidehop station, there are not too many cases where the same receiving frequency is used for all three directions, unless the guard angle between beam directions is at least 80 to 90 degrees. This is because of the wanted additional safety factor against the possibility of poor rejection by the adjacent antennas.

In the 6,700-mc band where passive reflectors are employed, it becomes very difficult to use the same frequency in even two directions which are 180 degrees apart. Back-to-back interferences at a repeater require separate frequencies for each direction, thus utilizing more of the frequency spectrum.

These and other similar interference problems will require considerable study for better co-ordination results. In congested areas, prior path testing in the different frequency bands and with different types of antennas may be necessary to assure the best possible use of the limited frequencies available.

Location Procurement and Accessibility

One of the problems about which little can be done is location availability and the economic factors of power and roads. A pattern of frequency assignments based on specific beamwidth patterns may be well planned to avoid interferences, but may not be possible to put into effect. Right of way and easements may be difficult or impossible to obtain for the planned locations.

This problem stresses the importance of initial engineering and independent field surveys to enable the potential user and the coordination group to make the best decisions as to frequency assignments prior to equipment purchase. It is also advisable that the potential user reserve his decision as to which commercial equipment to purchase until he has the answer to his land and location problem.

R-F Bandwidth, Frequency Stability, and Multiplexing Systems

The co-ordination problem here is a complex one because of the lack of standardization in system design and because of the several different commercial frequency bands available. Competitive forces, a large potential market, and the false feeling of no foreseeable congestion in the microwave bands will probably continue to keep alive these various differences.

Briefly, there are now available eight different types of systems, namely from General Electric, Motorola, Radio Corporation of America, Westinghouse, Federal, Collins, Philco, and Automatic Electric, with several other companies planning new systems soon to be released. Each of these systems has its relative advantages and disadvantages for particular applications. Until recently there has been little thought to co-ordination of different systems in the same area. As a result, a common allocation plan of r-f carrier frequencies and bandwidths will be difficult to determine unless some re-engineering of equipment is done.

An analysis of these various systems will show, however, that it may be practical to group them into bandwidth versus number of multiplex channels. For example, in a case where the ultimate need of telephone channels is only four, the systems authorized for licensing would be those requiring a minimum bandwidth for these four channels.

Following this line of reasoning, perhaps a division of systems into those most suitable for 5 mc, including the guard bands, and 10 mc, including the guard bands, would be a logical step towards the most efficient use of the spectrum bandwidth versus type of multiplex system, method of modulation, and frequency stability. This would allow assignment of a minimum required amount of bandwidth to qualified applicants and would permit them to engineer and purchase the system which makes the most efficient use of bandwidth.

User Co-ordination

A joint industry group called the Microwave Users Council has been formed recently. This council represents some 16 radio service groups in the category of safety and special services, such as the petroleum, power, railroad, forestry, police, highway, and other groups. The council will serve as a clearing house of information on installed and proposed non-common-carrier microwave communication systems. This information will be available to the prospective user and his engineer or consultant engaged in the planning of a new microwave system. For assurance that the type and routing of the microwave system planned is properly co-ordinated as to interference and future growth, it is advisable that an independent physical survey and study be performed to indicate the most suitable frequency band and the most appropriate microwave equipment for the application.

Reference

1. The Future of Microwaves, Edwin L. White. Petroleum Industry Electrical Association, Houston, Texas.

No Discussion

A Digital Data-Recorder for Dense Storage of Continuous Voltages

GERHARD L. HOLLANDER
ASSOCIATE MEMBER AIEE

Synopsis. The digital data-recorder stores in digital form continuous voltages from various transducers. It records for 8 minutes with 0.5-per-cent precision one channel of 1,600-cycle bandwidth and eight channels of 200-cycle bandwidth each. The design provides for additional amplitude-modulated (a-m) channels, with a bandwidth up to 80 kc if the operating time is reduced. The volume specification for the recorder is 500 cubic inches. Because of the digital representation of the data, more information can be stored in the limited space, the data can be reduced automatically, and calibration is simplified. While this paper gives block diagrams and wave forms for a specific application, the results are applicable for many instrumentation systems.

Introduction

THIS paper describes the design and operation of a magnetic digital data-recorder for torpedoes. To obtain certain characteristics of the submerged torpedo, a large amount of data must be recorded while the torpedo is under water. Similar recording requirements are encountered in the aircraft and missile fields. Existing recorders are inadequate for many applications where much information must be stored in limited space.

The information is recorded in digital form for compactness. An investigation showed that digital data representation uses the available space most economically for this application.[1]

Information recorded in digital representation is especially well suited for automatic data reduction. Labor savings of at least a factor of 100 can be realized. This is important, since the digital data-recorder stores so much more information than other recorders of the same size. However, digital representation does not preclude rapid qualitative examination of the records without complex data reduction equipment.

To store the data in digital form,[1] the information must be sampled. The channel bandwidth specifications listed here are in terms of samples per second. With approximately three samples per cycle of the highest information frequency in each channel, all information can be recovered with the specified accuracy.

SPECIFICATIONS

The minimum design-goal specifications of the digital data-recorder are shown in the following:

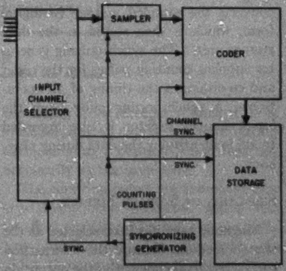


Fig. 1. Major units of recorder

Paper 54-318, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE-IAS-IRE-ISA Conference on Telemetering, Chicago, Ill., May 24–26, 1954. Manuscript submitted February 23, 1954; made available for printing May 12, 1954.

Gerhard L. Hollander is with the Clevite-Brush Development Company, Cleveland, Ohio.

This work was performed at the Servomechanisms Laboratory, Electrical Engineering Department, Massachusetts Institute of Technology, and was supported by the Bureau of Ordnance Contract *NOrd 11799* under the technical direction of the United States Naval Underwater Ordnance Station for the purpose of recording data from the performance of ranged torpedoes. The assistance of Thomas P. Sifferlen and the helpful criticism of Robert Kramer, both with the Massachusetts Institute of Technology, Cambridge, Mass., are gratefully acknowledged. This paper does not necessarily reflect the views of the Servomechanisms Laboratory or the Bureau of Ordnance.

Precision of the data	±0.5 per cent
Bandwidth of the precision channels:	
1 channel	5,000 samples per second
8 channels (each)	600 samples per second
Minimum operating time	8 minutes
Dimensions:	
Maximum mounting circle diameter	19 inches
Total volume	500 cubic inches
Ambient temperatures	0 to 66 degrees centigrade

In addition to the nine precision channels just mentioned, a-m channels with a bandwidth of 10,000 cycles per second are available. When the operating time is reduced proportionately, the bandwidth of the a-m channels can be increased to 80,000 cycles per second.

The nine precision channels can be combined into a single channel with a sampling rate of approximately 10,000 samples per second. Other combinations of the eight low-frequency channels can yield four channels with a sampling rate of 1,200 samples per second, or two channels of 2,400 samples per second. These changes do not require any internal adjustments of the recorder.

With the flexible building-block design, digital data-recorders can be packaged to other specifications as dictated by the application. Two typical alternatives are described later in the paper.

The Recorder

Essentially, the digital data-recorder consists of a 28-track magnetic tape recorder and auxiliary circuits that multiplex the input channels and convert the voltages into binary code; see Fig. 1. While the auxiliary circuits occupy as much space as the data storage unit, the more efficient data representation yields an over-all volume reduction for operating times of more than 5 minutes.

The major components of the recorder are shown in Fig. 1. The input-channel selector connects one of the input channels to the sampler, which samples the selected voltage quickly and holds it during the entire coding interval. The coder converts the sampled voltage to digital form, which is then stored in the data storage unit. The synchronizing generator supplies counting pulses for the coder and co-ordinates the timing of all operations. A synchronizing pulse from the input-channel selector is also recorded to aid in identifying the data during playback.

Selection of Storage Medium

Magnetic tape has been selected as the storage medium because the information can be recorded compactly, the output on playback is electrical, tape has a time dimension, and the techniques and components are well developed. Both magnetic tape and photographic film allow compact storage, but available digital transducers for high-speed photography are too large. Electric signals from the playback device are needed for automatic data reduction. With a time dimension in paper strips, magnetic tape, and photographic film, the location of the data indicates when they have been recorded without requiring elaborate programming.

Reason for Digital Data Representation

The advantages of digital representation for automatic data reduction have already been indicated. Now it will be shown that a-m and frequency-modulation (f-m) recorders could not meet the specifications to store the large amount of information in 500 cubic inches.

Amplitude modulation of magnetic flux density cannot meet the precision specification. Amplitude modulaton of spatial displacement, as exemplified by an oscillographic film recorder, can have sufficient resolution, but too much space may be needed. In an oscillographic film recorder the velocity of the film for 0.5-percent precision must be

$$v = 0.219f$$

where f is the highest signal frequency to be recorded with this precision. This means that for a 1,000-cycle-per-second signal the film speed must be at least 219 inches per second. Actually, a film speed of at least 440 inches per second should be used for a reasonable data reduction effort.[2] This requires more than 500 cubic inches for the space occupied by the film alone.

For f-m tape recording, rule of thumb indicates that for a 2-per-cent precision the carrier frequency should be at least 50 times the highest modulating frequency. According to preliminary results of another investigation, for 0.5-percent precision on a 1,000-cycle signal the carrier frequency should be at least 100 kc, requiring a tape speed of approximately 100 inches per second. With this speed, ten times faster than the normal tape speed in the digital recorder, the tape alone would occupy more than 500 cubic inches.

According to the specifications, in digital representation 10,000 samples per second must be stored for 500 seconds. In the digital data-recorder, the 2-inch tape stores 900 samples per inch, using a pulse density of 300 pulses per inch. This requires a tape speed of 11 inches per second. While the required tape volume is only 22 cubic inches, 120 cubic inches have been allowed for the tape and the necessary take-up and supply reels. Even with the 250 cubic inches of coding and multiplexing equipment, the 500-cubic-inch volume specification can be met.

Thus, in addition to the advantage of automatic data reduction, digital representation allows denser storage of the information. This yields a smaller over-all package if so much data must be recorded that the saving of storage space is greater than the volume occupied by the additional equipment which converts the information into digital code.

Ease of Calibration

When information is in digital form, its accuracy does not change with variations of circuit components. Thus only those circuits where the information is in analogue representation need to be calibrated. Since most of these analogue circuits are common to all input channels, usually only one channel needs to be calibrated. Furthermore, if a reference voltage like the transducer excitation is recorded in one channel, calibration is automatic many times each second.

Scope of Paper

The next section describes the operation of the recorder with block diagrams and wave forms. Then the permissible tolerances and errors for each component of the recorder are given. Finally, the important phase of data reduction is treated briefly.

Recorder Operation

This section describes how the digital data-recorder works. Fig. 2 is a detailed block diagram of the recorder. Dotted lines group the functional components as shown in Fig. 1. The timing schedule can be studied with the wave forms in Fig. 3. The recording of a single input channel will be discussed first, and the input multiplexing will be described later.

The *sampling repetition* period is the interval between two successive samples; in other words, it is the inverse of the sampling rate. The *sampling duration* is

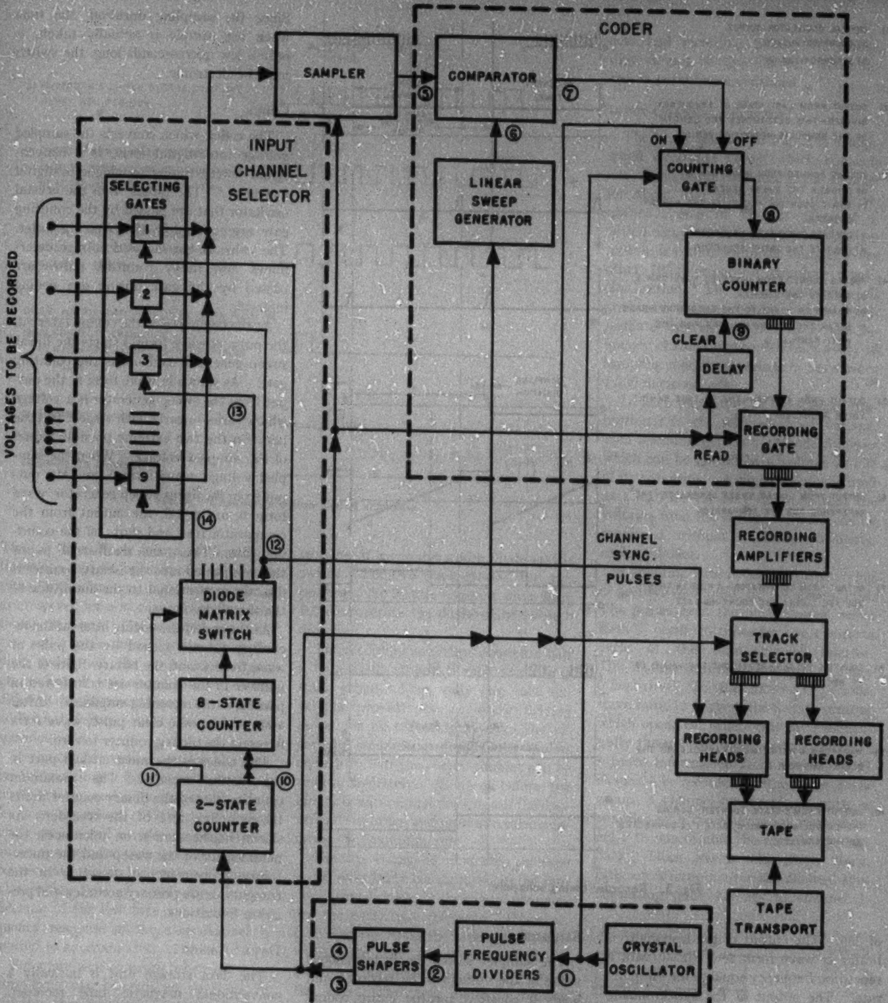

CODER
SAMPLER
COMPARATOR
⑤
⑦
⑥
ON
OFF
COUNTING GATE
LINEAR SWEEP GENERATOR
⑧
BINARY COUNTER
CLEAR
⑨
DELAY
READ
RECORDING GATE
INPUT CHANNEL SELECTOR
SELECTING GATES
VOLTAGES TO BE RECORDED
1
2
3
9
⑬
⑭
⑫
DIODE MATRIX SWITCH
8-STATE COUNTER
⑪
⑩
2-STATE COUNTER
CHANNEL SYNC. PULSES
RECORDING AMPLIFIERS
TRACK SELECTOR
RECORDING HEADS
RECORDING HEADS
TAPE
TAPE TRANSPORT
④
PULSE SHAPERS
②
③
PULSE FREQUENCY DIVIDERS
①
CRYSTAL OSCILLATOR
SYNCHRONIZING GENERATOR

Fig. 2. Block diagram of recorder

the brief interval when a voltage is actually sampled. Thus, the sampling repetition period is the sum of the *coding interval* and the *sampling duration and recovery time*.

For 0.5-per-cent over-all precision in the recorded data, about 200 counting pulses must occur during the coding interval. Additional time, between 20 and 50 counting pulses, must be allowed for the sampling duration and for the recovery time of the various circuits. Thus, a sampling repetition period 256 counting pulses long provides ample safety factor.

SYNCHRONIZING GENERATOR

The synchronizing generator provides the counting pulses for the coder and the synchronizing signals which co-ordinate the timing of the recorder. Each of these synchronizing signals occurs once during each sampling repetition period, but not necessarily at the same time.

The counting pulses are generated in a crystal-controlled oscillator which has excellent stability even without temperature stabilization. The frequency divider, three multivibrators, counts down the counting pulses by the chosen factor

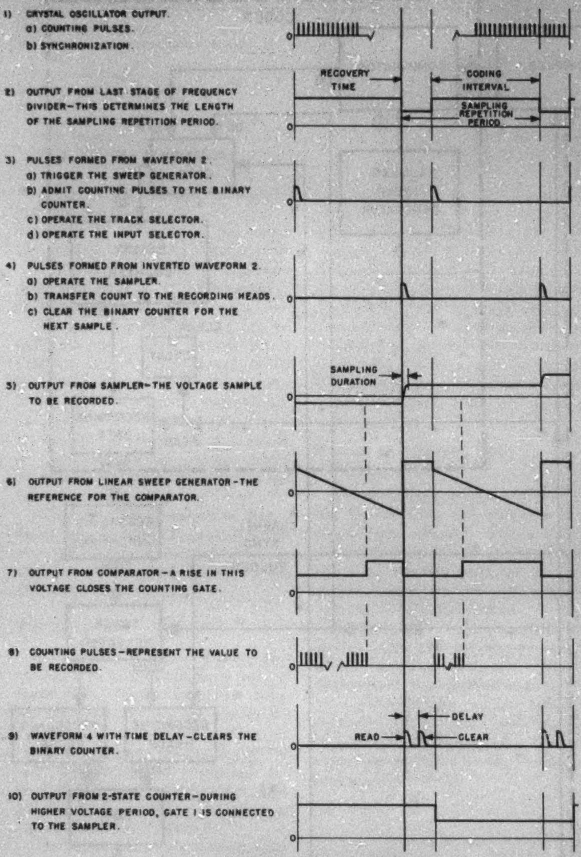

1) CRYSTAL OSCILLATOR OUTPUT.
a) COUNTING PULSES.
b) SYNCHRONIZATION.
2) OUTPUT FROM LAST STAGE OF FREQUENCY DIVIDER—THIS DETERMINES THE LENGTH OF THE SAMPLING REPETITION PERIOD.
3) PULSES FORMED FROM WAVEFORM 2.
a) TRIGGER THE SWEEP GENERATOR.
b) ADMIT COUNTING PULSES TO THE BINARY COUNTER.
c) OPERATE THE TRACK SELECTOR.
d) OPERATE THE INPUT SELECTOR.
4) PULSES FORMED FROM INVERTED WAVEFORM 2.
a) OPERATE THE SAMPLER.
b) TRANSFER COUNT TO THE RECORDING HEADS.
c) CLEAR THE BINARY COUNTER FOR THE NEXT SAMPLE.
5) OUTPUT FROM SAMPLER—THE VOLTAGE SAMPLE TO BE RECORDED.
6) OUTPUT FROM LINEAR SWEEP GENERATOR—THE REFERENCE FOR THE COMPARATOR.
7) OUTPUT FROM COMPARATOR—A RISE IN THIS VOLTAGE CLOSES THE COUNTING GATE.
8) COUNTING PULSES—REPRESENT THE VALUE TO BE RECORDED.
9) WAVEFORM 4 WITH TIME DELAY—CLEARS THE BINARY COUNTER.
10) OUTPUT FROM 2-STATE COUNTER—DURING HIGHER VOLTAGE PERIOD, GATE 1 IS CONNECTED TO THE SAMPLER.
RECOVERY TIME
CODING INTERVAL
SAMPLING REPETITION PERIOD
SAMPLING DURATION
DELAY
READ
CLEAR
0

Fig. 3. Recorder timing schedule

of 256. The output of the last multivibrator is wave form 2 of Fig. 3, with a repetition frequency equal to the sampling rate. The output of the synchronizing generator need only signify the beginning and the end of the coding interval. Because the coding interval takes up more than half of the sampling repetition period, the output-wave form 2 from the last multivibrator stage has been made unsymmetrical so that the positive portion determines the coding interval and the negative portion determines the sampling duration and recovery time. The synchronizing pulses shown in wave forms 3 and 4 are generated by simple resistance-capacitance pulse shapers.

Sampler

The sampler stores the instantaneous value of the voltage to be recorded and holds it constant for the entire coding interval. The coding interval is so long that without the sampler the input voltage would change considerably during this time. Such voltage change during the coding interval results in uncertainties which would preclude reconstitution of the signal with only three samples per cycle of the highest frequency.

Basically, the sampler consists of a storage capacitor and a single-pole single-throw switch, which connects the selected voltage to the capacitor and when the sampler is triggered by wave form 4. Since the sampling duration, the time when the sample is actually taken, is only a few microseconds long, the switch must be electronic.

Coder

The coder, which converts the sampled voltage into digital form, is a conventional sweep-timing analogue-to-digital converter.[3] The pulses from the crystal oscillator that are passed by the counting gate are counted in the binary counter. The value of the sampled voltage determines how many counting pulses are passed by the gate during one coding interval.

At the beginning of the coding interval, the pulse of wave form 3 starts the linear sweep generator and opens the counting gate. As shown in wave form 6, the output from the sweep generator is a voltage which varies linearly with respect to time between the two extreme possible values of the sampled voltage. When the sampled voltage, wave form 5, and the output from the linear sweep generator, wave form 6, are equal, the output from the comparator rises and shuts off the counting gate. Thus, the number of pulses that are passed into the binary counter is directly proportional to the amplitude of the sampled voltage.

At the end of the coding interval the recording gate is opened by the pulse of wave form 4, and the binary digits of the number in the counter are transferred in parallel to the recording amplifiers. After a short delay the clear pulse, wave form 9, resets the binary counter to zero.

The coder is the most critical part in the entire recorder. The maximum counting rate of the binary counter limits the sampling rate of the recorder. As shown in the section on tolerances, the nonlinearity of the sweep and the uncertainty of coincidence detection in the comparator are primary accuracy and precision limitations.

Data Storage

The data storage unit is basically a conventional magnetic tape recorder. An unusual feature is that successive samples (inputs) are stored in different track groups to utilize the full tape width. This section treats the components of the data storage unit and the arrangement for storing the digits of the information.

The present design provides for at least 28 tracks on the 2-inch magnetic tape. Since each coded sample contains *nine* binary digits, including one checking digit to be described later, the digits for each sample are recorded simultaneously in nine tracks, called a track group.

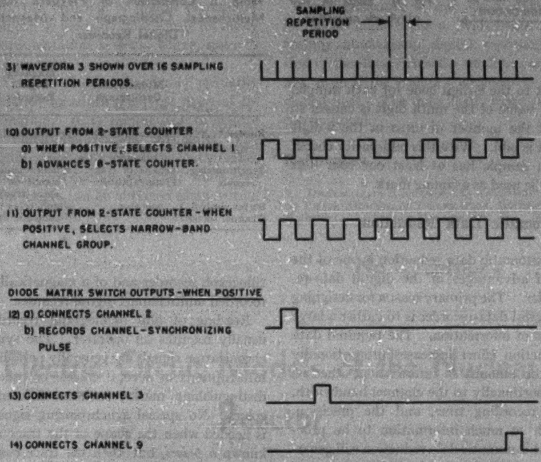

SAMPLING REPETITION PERIOD
3) WAVEFORM 3 SHOWN OVER 16 SAMPLING REPETITION PERIODS.
10) OUTPUT FROM 2-STATE COUNTER
a) WHEN POSITIVE, SELECTS CHANNEL 1
b) ADVANCES 8-STATE COUNTER.
11) OUTPUT FROM 2-STATE COUNTER - WHEN POSITIVE, SELECTS NARROW-BAND CHANNEL GROUP.
DIODE MATRIX SWITCH OUTPUTS - WHEN POSITIVE
12) a) CONNECTS CHANNEL 2
b) RECORDS CHANNEL-SYNCHRONIZING PULSE
13) CONNECTS CHANNEL 3
14) CONNECTS CHANNEL 9

Fig. 4. Channel-selection timing schedule

Two or three track groups of digital information can be accommodated on the tape, depending on how many tracks are reserved for a-m channels or for additional checking. For a given sampling rate and a given minimum spacing of the digit pulses on the tape, more track groups permit lower tape speeds. Thus, different arrangements provide different additional features over the minimum specifications.

For maximum recording time, the digital information is distributed into three track groups, and only one track is left for a-m recording. With three digital track groups and a pulse spacing of 300 pulses per inch, the tape speed is 11 inches per second. Seven-inch reels will provide a recording time of nearly 20 minutes. Because of the low tape speed, the frequency response of the a-m channel is limited to approximately 12 kc.

For more a-m channels when the running time can be reduced to 8 to 10 minutes, the digital information is divided into two track groups with a density of 250 pulses per inch. Since the tape now must operate at 20 inches per second, the frequency response of the a-m channels will be better than 20 kc.

Since digital information is recorded in only one track group at a time, one set of nine recording amplifiers can be shared by all track groups. The track selector is an electronic 9-pole, 2-position or 3-position switch that rotates the digits to the recording heads of the different track groups. In general, the track selector consists of a counter and a diode matrix switch. For two track groups, the output from the 2-state counter, wave form 10, can operate the diode matrix switch directly.

The recording amplifiers consist of half a twin-triode stage for each track in a track group. Thus, only five tube envelopes provide the necessary driving power for all recording heads. The recording amplifiers stretch the pulses enough so that the binary counter can be reset immediately. As stated before, the binary counter limits the speed of the digital data-recorder, and prompt resetting increases the sampling rate.

The tape transport must keep the tape speed only fairly constant, since the precisely timed pulses in the digital channels are the basic time reference. However, if during recording the tape speed varies widely, say over 10 per cent, pulses will be stored too densely in some sections of the tape. The a-m channels require much more constant speed. While the digital data pulses can be used to correct the tape speed during playback, this becomes increasingly difficult for speed variations above 1 per cent. Elaborate equipment for maintaining constant tape speed increases the size of the magnetic tape recorder. Therefore, a compromise is reached by using a motor with such torque-speed characteristics that provide as constant a tape speed as possible without special speed-regulating devices.

The a-m channels, not shown in Fig. 2, have a separate group of recording amplifiers and recording heads. Of course, they cannot be time-shared. Their design is entirely conventional.

Input-Channel Selector

The input selector in Fig. 2 is a high-speed electronic switch which connects one of the input channels to the sampler for about 100 microseconds in the sequence shown in Fig. 4. Channel 1, with a sampling rate of 5,000 samples per second, is connected to the sampler every other 100-microsecond sampling repetition period. The other eight channels are connected consecutively in the intermediate intervals. Thus, each low-frequency channel is recorded once every 16 sampling repetition periods or once every 1,600 microseconds.

To keep the output of the selector proportional to the selected input voltage, each selecting gate is a cathode follower which can be cut off by a voltage at one of its two grids. A single resistor is used as a load for the nine tubes, so that the voltages from the various channels can be separated without nonlinear elements, such as diodes.

A selecting gate is opened by applying a positive potential to its screen grid. The pulses from the pulse shaper, wave form 3, occur 16 times for one complete cycle of the input-channel selector. The 2-state counter opens gate 1 every other time. In the alternate periods, wave form 11 triggers the 8-state counter which opens the other eight gates sequentially through the diode matrix switch.

Since information from several input channels is recorded in the same track group, a channel-synchronizing pulse is recorded with the information from channel 2 to unscramble the records at playback. Thus, every odd (first, third, fifth, etc.) sample after the channel-synchronizing pulse belongs to channel 1.

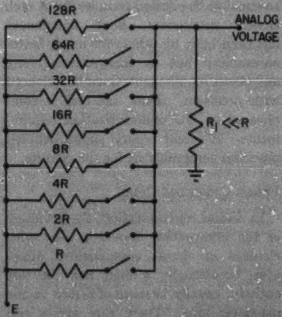


Fig. 5. A typical decoding scheme

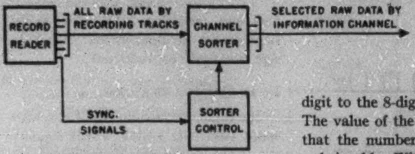

RECORD READER
ALL RAW DATA BY RECORDING TRACKS
CHANNEL SORTER
SELECTED RAW DATA BY INFORMATION CHANNEL
SYNC. SIGNALS
SORTER CONTROL

Fig. 6. Steps in any playback operation

Similarly, every second pulse after the synchronizing pulse belongs to channel 3; fourth pulse, to channel 4; etc.

Tolerances and Errors

The accuracy requirements of the digital data-recorder are very stringent. While the information is in analogue representation, the data are kept accurate by limiting the permissible deterioration in each stage. After the information has been quantized, it is accurate unless a digit is lost completely. This rare digit loss is detected by an error-checking code.

Tolerances of Analogue Components

In setting the tolerances of the analogue components, remember that the errors in the various components are statistically independent. Then, the total expected error is the square root of the sum of the squares of the expected component errors, the well-known rms relationship. The limits of signal deterioration allowed in each critical stage in per cent of the maximum range of the signal are:

Input-channel selector	±0.1
Synchronizing generator	±0.01
Sampler	±0.1
Linear sweep generator	±0.2
Quantizing error	±0.25
Comparator	±0.25

With this tolerance schedule, the maximum expected error is ±0.43 per cent.

These error values have been selected to equalize the design complexity of each critical component. Each component has been built by others for much better accuracies; but in this recorder, size and power requirements must be balanced with precision. Thus, the task is complicated by the miniaturization requirements. A recent report[4] shows detailed precision computations for the circuits.

Error-Checking Code

In digital representation, the accuracy of the information does not vary with changes in circuit components. However, occasionally a digit may be lost entirely, usually because of a flaw in the magnetic tape. These rare errors can usually be detected by adding a ninth digit to the 8-digit code for each sample. The value of the ninth digit is chosen so that the number of ONEs in the 9-digit code is odd. With this odd-parity check, each sample has at least one ONE that can be used as a timing mark.

Automatic Data Reduction

Automatic data reduction is one of the chief advantages of the digital data-recorder. The primary reason for designing a digital data-recorder is to gather a large mass of information. The required data reduction effort increases proportionally to the amount of information; that is, proportionally to the channel bandwidth, the recording time, and the precision. With so much information to be processed, automatic data reduction will generally cut the required labor by a factor of at least 100.

The complexity of the data reduction equipment depends on the desired computations and precision. For qualitative examination the digital data-recorder needs very little data reduction equipment. For precise computation of results, much time is saved with completely automatic data reduction using computers.

For minimum reduction equipment, eight relays can be connected as shown schematically in Fig. 5 to decode the digital data into analogue form. The output from the decoder is plotted on a pen recorder where it can be analyzed visually. This simplified arrangement loses the automatic-reduction advantage of the digital recorder, but it illustrates the flexibility of the reduction of digital data. Yet for precise computations, manual labor can be completely eliminated and results are available before they are obsolete.

Could other types of recording produce raw data suitable for fully automatic reduction? The main obstacle to automatic data reduction for most recording methods is the automatic reading and sorting of the information in the storage device.

A general block diagram of any playback device for a multichannel data system is shown in Fig. 6. The reader detects all signals in all channels. The synchronizing signals control a channel sorter which permits only the desired data channels to pass for further processing. Table I compares the components of the playback operations of a multichannel oscillograph recorder and of a magnetically recorded multichannel digital recorder.

Table I. Comparison of Playback from Multichannel Oscillograph and Magnetic Digital Recorder

	Multichannel Oscillograph	Magnetic Digital Recorder
Reader	operator or flying spot scanner	reading heads
Synchronizing signals	shape of trace interruptions	synchronizing signal on extra track
Sorter control Channel sorter	operator	simple relay or tube circuit

Reading of film oscillograph records usually requires an operator. The synchronization signals are generally periodic interruptions of several traces, or other distinguishing marks, such as different colors. No special synchronizing signal is needed when the shape of the trace is known *a priori*, but then the trace contains less information. If each trace were confined to a separate portion of the film, a flying spot scanner can reduce the records automatically; but this also limits the total excursion of each trace to a fraction of the film, thus lowering the resolution.

In the case of the digital data-recorder, the magnetic tape is read by standard magnetic reading heads. The synchronizing signals, which are recorded in a separate track, operate the channel sorter. The channel sorter and sorter control are essentially an inverse arrangement of the selecting gates in the digital data-recorder. Since playback can proceed much more slowly than recording, relays can be used in place of the diodes and tubes in the recorder, thus insuring reliability and minimum voltage drop across the gates.

While automatic playback of digitally recorded data presents no difficulty, most other common recording methods require a human operator when the storage medium is used efficiently.

Conclusions

This paper has presented the design of a recorder that stores in digital form continuous voltages from various transducers. In digital representation:

1. More information can be stored in limited space.
2. The data can be reduced automatically.
3. Calibration is simplified, because the same critical circuits are shared by all channels.

While this paper describes a specific design, the principle of digital representation can be used advantageously for many instrumentation systems where space is at a premium.

The digital data-recorder is primarily intended for storing a large mass of information. It is impractical to reduce so much information manually. The digital representation of the data makes automatic reduction possible; while analogue traces can be obtained easily with a few relays and a pen recorder.

The significance of this design is that for most high-precision, wide-band recording application, digital representation of the data yields the smallest package, even if the coding equipment is included in it. The individual circuits are conventional.

References

1. DIGITAL DATA RECORDER—AN INVESTIGATION OF DENSE STORAGE OF INFORMATION, Gerhard L. Hollander. *Technical Memorandum No. 6897-TM-12*, Massachusetts Institute of Technology, Servomechanisms Laboratory, Cambridge, Mass., 1953, pp. 47–54.

2. *Ibid.*, p. 25b.

3. CRITERIA FOR THE SELECTION OF ANALOG-TO-DIGITAL CONVERTERS, Gerhard L. Hollander. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 9, 1953, p. 673.

4. DESIGN STUDY OF A DIGITAL DATA-RECORDER, Gerhard L. Hollander, Thomas P. Sifferlen. *Engineering Report No. 6897-ER-6*, Massachusetts Institute of Technology, Servomechanisms, Laboratory, Cambridge, Mass., April 7, 1954.

5. BIBLIOGRAPHY ON DATA STORAGE AND RECORDING, Gerhard L. Hollander. *AIEE Transactions*, vol. 73, pt. I, March 1954, pp. 49–58.

No Discussion

Electric Circuit Models of the Nuclear Reactor

GABRIEL KRON
FELLOW AIEE

Synopsis: A series of equivalent circuits are established for the partial differential equations representing the diffusion of neutrons in a reactor. One-, two-, and three-dimensional spatial models, valid for all curvilinear orthogonal reference frames, are given first for monoenergetic neutrons. The presence of variable-energy neutrons is represented next by the appearance of a fourth dimension. Variation of neutron density in time is also considered along an extra fifth dimension. Circuit models with arbitrary shapes and with tens of thousands of net points may now be solved quickly and efficiently with the aid of digital computers by the recently developed tensorial method of tearing the model apart into smaller parts and solving each part separately. The effect of movable and variable control fixtures in the reactor, also a variation in composition, may now be considered without repeating the solutions arrived at for the rest of the reactor.

Introduction

Electric Power Engineers and Electronics

AS THE first large-scale industrial application of nuclear energy appears to be at present the generation of electric power, the electric power engineer is again being pressed by circumstances outside his control to get acquainted with a new physical science, namely, nucleonics and quantum mechanics. Three decades ago a similar situation confronted the electric power engineer when the arrival of radio forced him to study electronics and electromagnetic wave theory.

No one can deny that the sciences of waves and electronics both profited by concepts of the electric power engineer, namely, by the concepts of impedance, of equivalent circuit etc., after suitable generalizations, of course. There is every reason to believe that in the coming coalition between the electric power engineer and nuclear engineer the latter will also profit by his contact with the power engineer and his basic concepts.

Electric Power Engineers and Nucleonics

One such contact exists already in the form of an electric circuit model for the wave equations of Schrödinger.[1] These equations play the same basic role in quantum mechanics and nucleonics that the field equations of Maxwell do in radio and electronics. The familiar concepts of natural frequencies of an electric circuit and the distribution of potentials in a circuit may replace the eigen-value and eigen-function concepts of the physicist, customarily wrapped in the language of Hermite and other polynomials.

The present paper establishes another contact between the electric power engineer and the nuclear engineer. A set of electric circuit models (equivalent circuits) are established for the linear partial differential equations representing the diffusion of neutrons in a nuclear reactor of arbitrary shape. The well-known equations of diffusions may be expressed in any arbitrary curvilinear co-ordinate system along one, two, or three spatial dimensions. The variable energy of neutrons is represented in the model either more exactly as a fourth dimension, or approximately as a parameter of the multigroup theory. The variation of neutrons in time is introduced also as an additional fifth dimension. Of course, any combination of spatial, energy, and time dimensions may be assumed as desired.

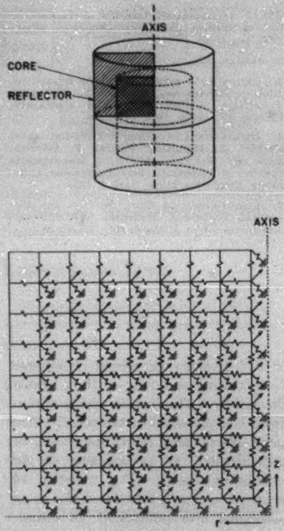
Nuclear Reactor as a Transmission System

The electric circuit model of a reactor is analogous to an electric power transmission network and it gives a visual picture of several of the processes (greatly simplified) that take place in a reactor. Each generator (impressed junction current) represents a generation of neutrons by fission. The loads (impedance to ground) stand for the absorption of neutrons by the materials in the reactor, while the flow of electric current along the transmission network itself represents the diffusion of neutrons (the neutron current) from point to point within the reactor.

Paper 54-199, recommended by the AIEE Nucleonics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted August 11, 1953; made available for printing March 22, 1954.

GABRIEL KRON is with the General Electric Company, Schenectady, N. Y.

The work described in this paper was done for the Knolls Atomic Power Laboratory operated for the U. S. Atomic Energy Commission by the General Electric Company, Contract *No. W-31-109, Eng. 52.* This project is part of the general program being conducted by the Nuclear Engineering Unit of the Laboratory, supervised by Harry E. Stevens, Jr. The author wishes to acknowledge his indebtedness to Drs. Leon T. Anderson and Martin Storm for many conversations and suggestions concerning the general problem. Acknowledgements are also due to Drs. L. Tonks, R. Ehrlich, and H. Hurwitz of the Laboratory for their help in the 1-dimensional (spherical) circuit studies.



$$-\frac{\partial}{\partial r}\left(a_r r\frac{\Delta z}{\Delta r}\right)\frac{\partial}{\partial r}\Delta r^2\phi-\frac{\partial}{\partial z}\left(a_z r\frac{\Delta r}{\Delta z}\right)\frac{\partial}{\partial z}\Delta z^2\phi+ b(r\Delta r\Delta z)\phi=c(r\Delta r\Delta z)$$

Fig. 1. Cylindrical reactor with core and reflector

The absolute potential at every junction is equal to the neutrons of some particular energy appearing at that point of space.

Several of the equivalent circuits were published by the author in 1945,[2] without, however, any reference to application to the nuclear reactor. The present paper merely reinterprets the circuit symbolism in terms of the declassified and published neutron diffusion equations.[3] The paper also assumes special and more general cases.

Also it may be mentioned that the 1-dimensional circuit model has been used successfully for the last 6 years for the analytical calculation of various types of nuclear reactors with spherical symmetry. The two and higher dimensional circuits, however, could not be put to use until recently for two reasons:

1. The inadequacy and inexactness of all existing d-c and a-c network analyzers.
2. The lack of any analytical method to solve extensive 2-dimensional electric networks quickly and in an organized manner.

Solving Large Physical Systems in Easy Stages

Recently the author succeeded in developing a highly efficient method of solving physical systems in general and electric networks in particular—possessing many thousands of variables—by means of already available high-speed digital computers. The new method consists of tearing the given network (or physical system) apart into several smaller networks and solving each subdivision separately. The component solutions may afterward be interconnected to give the solution of the original large network. The remaining work consists of solving for the comparatively few number of variables appearing at the cuts. The latter solution may be accomplished (if necessary) by further tearing apart the cuts themselves.

This method of interconnecting solutions has already been applied to three different types of electric networks; namely to:

1. Electric power transmission networks[4-7] in which the known variables are not currents **i** but the far more difficult concepts of real powers **P**.
2. The partial differential equations of Poisson and Laplace.[8]
3. The 2-dimensional field equations of Maxwell.[9]

Singly and Multiply Grounded Networks

It should be pointed out that, of these three types of circuit configurations, the solution of the Maxwellian circuits by tearing involves the least number of new concepts and is the fastest, because of the existence of ground impedances at every junction point. All the equivalent circuits of the nuclear reactor have the same configuration as those of the 2-dimensional field equations of Maxwell (with pure resistances, though). Hence, the simplified method of solution of the equation **I=YE**, outlined in reference 9 for the field equations of Maxwell, are also valid without any change to the solution of the diffusion equations of the reactor **I=YE** if the impressed currents **I** are assumed to be known and the junction potentials **E** are assumed to be unknown.

The boundary conditions in the nuclear reactor are also either open circuits, short circuits, or known impedances, etc. The problem of representation of free space extending to infinity arises in the reactor only as that of a short, extrapolated distance.

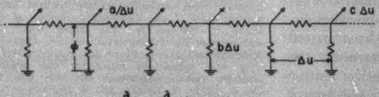
Reactors with Variable or Movable Parts

The electric circuit model of **I=YE** offers either new types of iterative procedures to find **E** (or **Z**) for a given **I** or exact procedures to find **Z** outright.

If the impedance matrix **Z** of the reactor is calculated (instead of **E** only), it is possible to save a large amount of labor in cases where it is necessary to make changes in portions of the reactor. If the composition of certain component parts is changed, then it is not necessary to repeat the **Z** calculation for the rest of the reactor. Only the **Z** of the changed portion of the reactor needs to be recalculated. The new component **Z** is then interconnected with the permanent **Z** of the rest of the reactor. If the angular or linear positions of some component parts are changed for control purposes, then only the interconnection of their **Z** need be changed, leaving all component **Z**'s themselves unchanged.

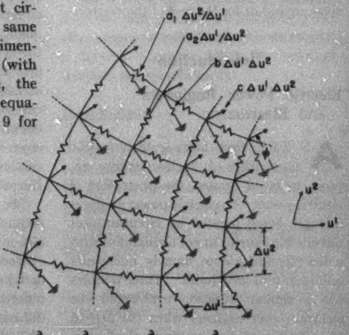
The impedance matrices of cores, reflectors, and blankets with a large variety of compositions may thus be interconnected with a comparatively small amount of labor. Control problems involving movable or variable fixtures be-

Fig. 2 (below). One-dimensional network for one energy group

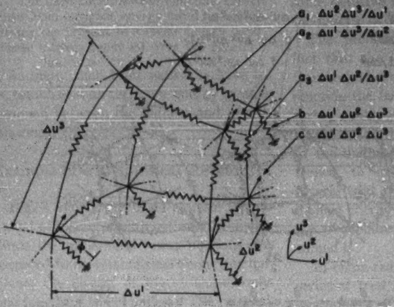


$$-\frac{\partial}{\partial u}a\frac{\partial}{\partial u}\phi+b\phi=c$$

Fig. 3 (right). Two-dimensional network for one energy group

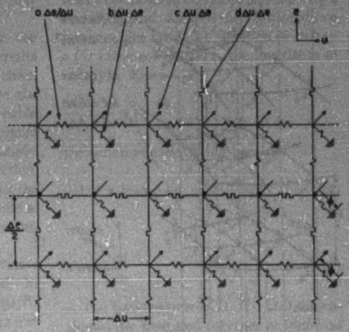


$$-\frac{\partial}{\partial u^1}a_1\frac{\partial}{\partial u^1}\phi-\frac{\partial}{\partial u^2}a_2\frac{\partial}{\partial u^2}\phi+b\phi=c$$


$a_1\ \Delta u^2 \Delta u^3/\Delta u^1$
$a_2\ \Delta u^1 \Delta u^3/\Delta u^2$
$a_3\ \Delta u^1 \Delta u^2/\Delta u^3$
$b\ \Delta u^1 \Delta u^2 \Delta u^3$
$c\ \Delta u^1 \Delta u^2 \Delta u^3$
Δu^3
Δu^2
Δu^1
ϕ
u^3
u^2
u^1

$$-\frac{\partial}{\partial u^1} a_1 \frac{\partial}{\partial u^1} \phi - \frac{\partial}{\partial u^2} a_2 \frac{\partial}{\partial u^2} \phi - \frac{\partial}{\partial u^3} a_3 \frac{\partial}{\partial u^3} \phi + b\phi = c$$

Fig. 4 (left). Three-dimensional network for one energy group



$$-\frac{\partial}{\partial u} a \frac{\partial}{\partial u} \phi + b\phi + d \frac{\partial \phi}{\partial e} = c$$

Fig. 5 (right). One-dimensional reactor with all energy groups

come thereby amenable to more exact and still not too lengthy analysis.

Advantages of the Method of Tearing

The tensorial method of tearing a large system into several component parts and interconnecting the component solutions shows up to advantage not only when movable or variable parts exist in the reactor. Since a nuclear reactor usually can be subdivided in such a manner that several of the parts are identical, the solution of one part represents the solution of all the identical parts. The following additional advantages may also be claimed for solving nuclear reactor problems by the method of tearing:[4,9]

1. Existing digital computers may be used to solve reactors with thousands of assumed points.
2. The inversion time (calculation of **Z** from a given **Y**) is reduced to a fraction of about $2/n^2$ of conventional inverse calculation time, where n is the number of subdivisions. (This rule of thumb assumes that a specialized method is not used for the conventional inverse calculations, only the Gauss-Crout elimination method.)
3. The final inverse matrix **Z** comes out in a factorized form, containing only a small fraction (about $1/\sqrt{n}$) of nonzero elements of the conventional inverse.
4. The accuracy of calculations depends on the size of the subdivisions and the nature of cuts. The number of significant figures retained no longer depends on the size of the original system.
5. The correctness of coding, of computer calculations, or of the analytical method itself may be checked at almost every step by subjecting the known results to physical tests (Kirchhoff's laws, etc.).
6. The calculation of **Z** consists of a succession of addition, multiplication, and division of matrices. Hence, no special detailed subroutine coding of the computers is needed, only a sequence coding of matric manipulations.
7. The equations of the given original systems are never written down, only those of subdivisions. Identical subdivisions are recorded only once, even though a large number of them may be identical.
8. Since a large part of the final factorized inverse matrix has many identical matrices, only a fraction of the factorized inverse need be stored for the calculation of **E** for an assumed **I**.

Steady-State, Monoenergetic Equivalent Circuits

Analogies Between Reactors and Networks

In an actual electric network the electric charges are bodily transported from one portion of the network to another. When a network represents a nuclear reactor, then the motion of electric charges represents the motion of neutrons. The three basic concepts of a network, namely voltage e, current i, and impedance z (or admittance y) correspond to the following three basic concepts of the reactor:

1. The absolute potential of a junction E represents the "neutron flux" nv, where n is the number of neutrons and v their velocity. That is, E represents the number of neutrons that are available to take part in the various reactions.
2. The impedance z (or admittance y) of a coil represents the cross section Σ of a point of the reactor with respect to some particular nuclear reaction. (The inverse of Σ is the mean free path λ of a particular reaction.) In diffusion (transport) phenomena the concept of a cross section represents an impedance z, in absorption and scattering phenomena an admittance y.
3. The current i in a coil represents the neutron current I flowing at the same point of the reactor in the same direction. It represents the flow of neutrons taking part in a particular reaction.

Ohm's law for each reaction is then

$$i = ye \,|\, I = \sum (nv) \qquad (1)$$

It is interesting that the coils, representing the cross sections (z or y) to different nuclear reactions, are orthogonal to each other in the model. That is, all types of reactions, such as diffusion, absorption, scattering, or fission, are represented by spatially or topolcgically orthogonal currents.

Conditions of Continuity

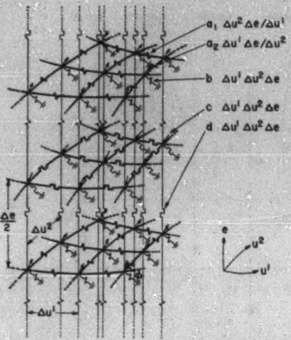
It should be emphasized that it is possible to assume other correspondence between circuit and reactor quantities. For instance, the absolute potential E may represent slowing-down density q, or number of neutrons n. However, this correspondence has the great advantage that at the interfaces of two different materials both the neutron flux nv and the neutron current I are continuous, as are the analogous voltages E and currents i.

Hence, with the use of neutron flux nv as voltage E no transformation of reference axes need be introduced between two different materials and the composition of the reactor may vary from point to point in an arbitrary manner.

The Conservation of Neutrons

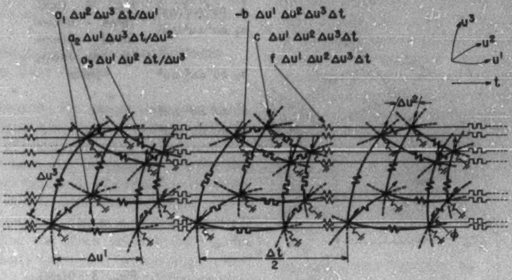
In nuclear reactor calculations it is necessary to account for all neutrons from their instant of birth (by fission) to their final expiration by leakage or by absorption. That is, for each part of the reactor (core or reflector, etc.) or for the entire reactor at all instants, the following equalities must be satisfied

$$\sum \text{source} = \sum \text{absorption} + \sum \text{leakage}$$


$a_1\ \Delta u^2 \Delta e/\Delta u^1$
$a_2\ \Delta u^1 \Delta e/\Delta u^2$
$b\ \Delta u^1 \Delta u^2 \Delta e$
$c\ \Delta u^1 \Delta u^2 \Delta e$
$d\ \Delta u^1 \Delta u^2 \Delta e$
$\frac{\Delta e}{2}$
Δu^2
e
u^2
u^1
Δu^1

$$-\frac{\partial}{\partial u^1} a_1 \frac{\partial}{\partial u^1}\phi - \frac{\partial}{\partial u^2} a_2 \frac{\partial}{\partial u^2}\phi + b\phi + d\frac{\partial \phi}{\partial e} = c$$

Fig. 6. Two-dimensional reactor with all energy groups


$a_1\ \Delta u^2 \Delta u^3 \Delta t/\Delta u^1$
$a_2\ \Delta u^1 \Delta u^3 \Delta t/\Delta u^2$
$a_3\ \Delta u^1 \Delta u^2 \Delta t/\Delta u^3$
$-b\ \Delta u^1 \Delta u^2 \Delta u^3 \Delta t$
$c\ \Delta u^1 \Delta u^2 \Delta u^3 \Delta t$
$f\ \Delta u^1 \Delta u^2 \Delta u^3 \Delta t$
u^3
u^2
u^1
t
Δu^2
Δu^3
Δu^1
$\frac{\Delta t}{2}$
ϕ

$$-\frac{\partial}{\partial u^1} a_1 \frac{\partial}{\partial u^1}\phi - \frac{\partial}{\partial u^2} a_2 \frac{\partial}{\partial u^2}\phi - \frac{\partial}{\partial u^3} a_3 \frac{\partial}{\partial u^3}\phi + b\phi + f\frac{\partial \phi}{\partial t} = c$$

Fig. 7. Three-dimensional reactor with all energy groups

It is interesting that in an electric circuit model the conservation of neutrons is exactly satisfied at all instants because of Kirchhoff's laws. In Fig. 1 the summation of impressed currents (source neutrons) is exactly equal to the summations of currents entering the grounded conductors (absorbed neutrons) plus the summation of currents entering the boundary short circuits (neutrons leaking out of the reactor).

Orthogonal, Curvilinear Reference Frames

All equivalent circuits in this paper are expressed along orthogonal curvilinear co-ordinates. Every symbol stands for a line integral, surface integral, or volume integral of some quantity.[10–12] In particular the following tensor densities are introduced where the primed quantities are arbitary functions of the variables u^1, u^2, u^3.

$$a_1 = \frac{h_2 h_3}{h_1} a_1'; \; a_2 = \frac{h_1 h_3}{h_2} a_2'; \; a_3 = \frac{h_1 h_2}{h_3} a_3'$$

$$b = h_1 h_2 h_3 b'; \; c = h_1 h_2 h_3 c'; \; d = h_1 h_2 h_3 d' \qquad (2)$$

Also each block (with length Δu^α) may have arbitrary unequal length along any direction. As in all difference representations of a differential equation, a half-block inaccuracy exists along all dimensions considered.

The values of the variables u^α and the metrical coefficients h are given in Table I for the three most common orthogonal reference frames. Since the tensor densities of equation 2 represent surface and volume integrals, their extension to arbitrary reference frames and to arbitrary shape and size of blocks are self-evident.

Steady-State, Monoenergetic Models

All more general circuit models are built up of a series of basic models. The basic model itself implies two special assumptions:

1. The neutron flow is steady. That is all time variations are ignored temporarily.
2. All neutrons have the same energy, or rather, occupy the same narrow energy band. That is, for each energy group a different network is established.

The basic model represents some instantaneous spatial distribution of neutrons having the same energy. Accordingly the basic model itself may be a 1-, 2-, or 3-dimensional network, each dimension representing a spatial axis, such as x,y,z, or r,θ,z, etc. The basic model also represents the "multigroup" theory of the reactor in which the spatial distribution of neutrons belonging to a series of energies (or bands of energies) is studied, one energy at a time, in descending order from a higher to a lower energy band. That is, as many separate networks are established as there are energy groups assumed.

The more general circuits, involving variation in energy and in time (to be developed later on), consist of a series of such basic circuits, each basic circuit representing a spatial distribution of neutrons with different energies and at different instants.

Table I. Values of u^α and h

	u^1	u^2	u^3	h_1	h_2	h_3
Cartesian	x	y	z	1	1	1
Cylindrical	r	θ	z	1	r	1
Spherical	r	θ	ψ	1	r	$r \sin\theta$

The Basic Model

The steady motion of monoenergetic neutrons in a reactor is analogous to the diffusion of heat in a nonhomogeneous body, with the additional complication that not only sources but also absorption of neutrons exist at every point of space. Accordingly three sets of constants occur in the diffusion equation

$$-\text{Div}\ \mathbf{a}\ \text{grad}\ \phi + b\phi = c$$

$$-\nabla\ \mathbf{a}\ \nabla\phi + b\phi = c$$

$$-\nabla \mathbf{y}_a \nabla E + y_a E = I \qquad (3)$$

1. The Laplacian networks of Figs. 2–4 show the manner of diffusion of neutrons throughout the entire reactor. The coils are pure resistances, $1/a$, representing the transport cross sections Σ_{tr}. The absolute potential E of each junction represents the neutron flux $\phi = nv$. The current in each coil represents the neutron-current vector **i**

$$\mathbf{i} = \mathbf{y}_d\ \text{grad}\ \mathbf{E} \qquad (4)$$

Each admittance y_d and y_a are arbitrary functions of the space variables u^1, u^2, and u^3.

2. From each junction point an admittance b leads to ground, forming paths for the absorption currents. (The neutrons are captured by both fissionable and nonfissionable materials at every point of the reactor.)

3. The currents I impressed at the junctions represent the neutron sources. In the basic model two such sources may exist:

a. Neutrons produced by fission. (These occur only in the core materials.)

b. The difference between neutrons slowed down from a higher energy band and those leaving the present to a lower energy band. (These sources occur all over the reactor.)

As many such basic networks have to be established as there are energy groups assumed between the highest and lowest (thermal) energies.

Example of a Cylindrical Reactor

Fig. 4 shows the basic electric circuit model of a reactor with cylindrical shape. Since the upper and lower portions are assumed to be symmetrical about a center plane, the circuit of Fig. 1 shows only one-fourth of the cylinder. The internal core is assumed to be also of cylindrical shape. The core is shown to be surrounded by a reflector.

The outer or extrapolated boundaries (at a small distance outside the boundaries) appear in the model short-circuited to the ground since no neutrons are assumed to exist there. The flow of neutrons across the axis and across the central plane is zero; hence, the circuit model is open-circuited at those internal points.

Two sets of impressed currents **I** exist. In particular the fission currents $\mathbf{I}^f$ are impressed in the core and the slowing-down currents $\mathbf{I}^s$ in both core and reflector.

The Chain-Reaction Problem

The problem in each network may be formulated as follows. Given the equation of the network **I**=**YE**, where **Y** and **I** are known; find the vector **E**. The latter may be found either outright by some iterative method or by finding first the inverse of **Y**, namely **Z**.

The problem is complicated by the fact that the impressed fission currents (sources) $\mathbf{I}^s$ are not known. In a chain reaction the latter are related to the absorbed neutrons $\mathbf{I}^a$ captured by fissionable material, and the latter in turn are related to the unknown **E** of all energy bands. Hence, to maintain a chain reaction, several trial $\bar{\mathbf{I}}^f$'s must be assumed in the highest energy band until the calculated **E**'s of all the energies balance the assumed $\mathbf{I}^f$.

In chain-reaction calculations, when a series of trial $\mathbf{I}^f$'s must be assumed, the availability of the permanent **Z** matrix for each energy group enables the immediate calculation of the vector **E** by **ZI**.

More General Equivalent Circuits

Slowing Down of Neutrons

In the multigroup theory of the reactor as many basic networks are constructed as there are energy bands assumed between the highest and lowest energies. No electric connections exist between the basic networks. The number of higher energy neutrons entering the lower energy neutrons leaving each band (or group) are estimated for every point of space. An approximate method of estimation is given further on.

The partial differential equation representing also the flow of neutrons between the various energy groups is

$$-\text{Div } a \text{ grad } \phi+b\phi+\frac{\partial d\phi}{\partial e}=c \tag{5}$$

where the product $d\phi$ is the slowing-down density q and d itself the slowing-down power. The latter is a function of energy e (or rather of lethargy, the logarithm of energy). Since

$$\frac{\partial d\phi}{\partial e}=\frac{\partial d}{\partial e}\phi+d\frac{\partial \phi}{\partial e}$$

the differential equation becomes

$$-\text{Div } a \text{ grad } \phi+b'\phi+d\frac{\partial \phi}{\partial e}=c \tag{6}$$

where

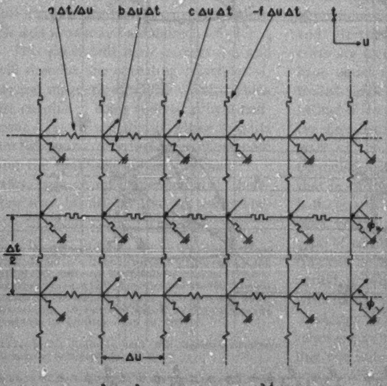
$$b'=b+\frac{\partial d}{\partial e} \tag{7}$$

The basic networks (1-, 2-, or 3-dimensional) are now tied together at identical space points by a set of alternating plus or minus resistances $1/d$, along which the neutron currents between the various energy groups flow; see Figs. 5–7. The basic networks also assume alternating plus or minus resistances and may be considered to extend along a fourth dimension e (lethargy).

Multigroup Theory

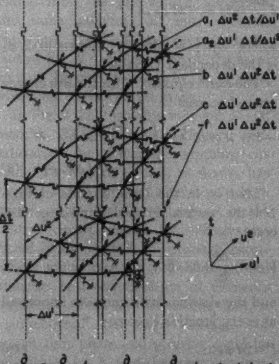
Equation 6 suggests an approximate method of finding the slowing-down currents $\mathbf{I}^s$ to be impressed at every junction of an electrically isolated basic network of Figs. 1–4, used in the multigroup theory.

Let the basic networks of Figs. 1–4 represent the nth energy group and let it be assumed that the neutron flux ϕ_{n-1} of the $n-1$ (the next highest) energy group has already been calculated. Then in the exact equation 6 it may be assumed as a first approximation that



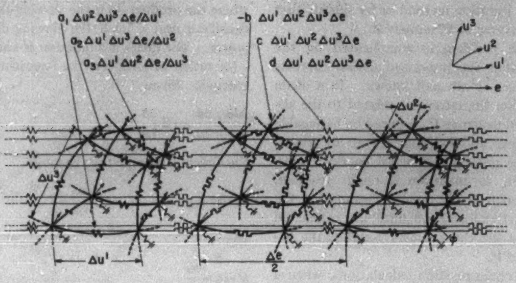
$$-\frac{\partial}{\partial u}a\frac{\partial}{\partial u}\phi+b\phi+f\frac{\partial\phi}{\partial t}=c$$

Fig. 8 (left). Time variation of a 1-dimensional reactor



$$-\frac{\partial}{\partial u^1}a_1\frac{\partial}{\partial u^1}\phi-\frac{\partial}{\partial u^2}a_2\frac{\partial}{\partial u^2}\phi+b\phi+f\frac{\partial\phi}{\partial t}=c$$

Fig. 9 (right). Time variation of a 2-dimensional reactor


$a_1\,\Delta u^2\Delta u^3\Delta e/\Delta u^1$
$a_2\,\Delta u^1\Delta u^3\Delta e/\Delta u^2$
$a_3\,\Delta u^1\Delta u^2\Delta e/\Delta u^3$
$-b\,\Delta u^1\Delta u^2\Delta u^3\Delta e$
$c\,\Delta u^1\Delta u^2\Delta u^3\Delta e$
$d\,\Delta u^1\Delta u^2\Delta u^3\Delta e$
u^3
u^2
u^1
e
Δu^2
Δu^3
Δu^1
$\frac{\Delta e}{2}$
ϕ

$$-\frac{\partial}{\partial u^1}a_1\frac{\partial}{\partial u^1}\phi-\frac{\partial}{\partial u^2}a_2\frac{\partial}{\partial u^2}\phi-\frac{\partial}{\partial u^3}a_3\frac{\partial}{\partial u^3}\phi+b\phi+d\frac{\partial\phi}{\partial e}=c$$

Fig. 10. Time variation of a 3-dimensional reactor

(ignoring for the present the important questions of average and point values)

$$\frac{\partial\phi_n}{\partial e_n}=\phi_{n+1}-\phi_{n-1}$$

where ϕ_{n+1} is unknown. Let it be assumed further, again as a first approximation, that (for the purpose of eliminating ϕ_n)

$$\phi_n=\frac{\phi_{n+1}+\phi_{n-1}}{2}$$

The unknown ϕ_{n+1} is then

$$\phi_{n+1}=2\phi_n-\phi_{n-1}$$

(ϕ_n is average, but ϕ_{n+1} and ϕ_{n-1} are not averages but point values.) Hence, as a first approximation, in the nth energy group the partial derivative may be replaced by

$$\frac{\partial\phi_n}{\partial e_n}=2\phi_n-2\phi_{n-1}$$

$$d\,\frac{\partial\phi}{\partial e}=2d\,\phi_n-2d\,\phi_{n-1}$$

where ϕ_{n-1} is already known and ϕ_n is the dependent variable.

The exact equation 3 for the nth energy group becomes

$$-\text{Div}\ \mathbf{a}\ \text{grad}\ \phi_n+\left(b+\frac{\partial d}{\partial e}+2d\right)\phi_n=2d\,\phi_{n-1}+c \qquad (8)$$

That is, in the basic circuits of Figs. 1–4 the vertical ground impedances become

$$b'=b+\frac{\partial d}{\partial e}+2d \qquad (9)$$

and the slowing-down currents impressed at every junction become

$$I^s=2d\,\phi_{n-1} \qquad (10)$$

where ϕ_{n-1} is known from the previous (next higher energy) circuit calculations. The quantity c represents the fission current that may exist at the core junctions. (The c's also are assumed to be known.) More exact approximations are also possible.

Time Variation of Neutrons

Of the many possible time delays entering into the various processes, only one will be considered, namely, the time of diffusion of neutrons. Assuming monoenergetic neutrons, equation 3 is generalized to

$$-\text{Div}\ \mathbf{a}\ \text{grad}\ \phi+b\phi+f\frac{\partial\phi}{\partial t}=c \qquad (11)$$

This equation looks the same as equation 5 representing the variation (slowing down) of neutrons along the energy scale. That is, energy (or rather lethargy) is now replaced by time. Accordingly, the equivalent circuits of Figs. 8–10 look the same as Figs. 5–7, except that the fourth dimension is now time t instead of lethargy e.

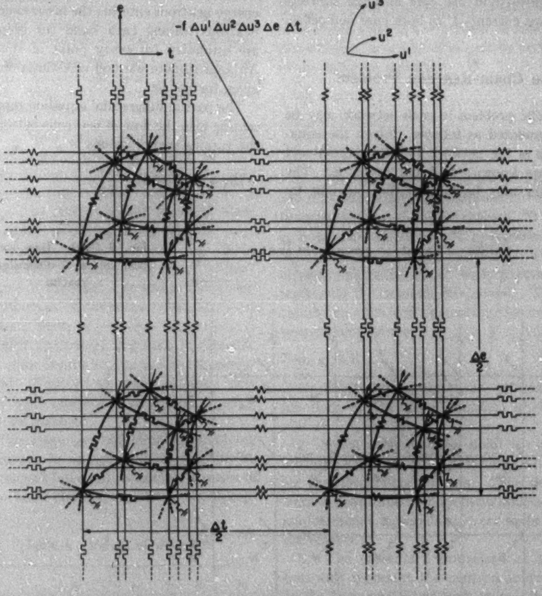
If the slowing down of neutrons is also considered, equation 6 is generalized to

$$-\text{Div}\ \mathbf{a}\ \text{grad}\ \phi+b'\phi+d\frac{\partial\phi}{\partial e}+f\frac{\partial\phi}{\partial t}=c \qquad (12)$$

The corresponding 5-dimensional circuit is shown in Fig. 11. A set of 3-dimensional basic circuits are arranged in a 2-dimensional array along energy and time.

References

1. ELECTRIC CIRCUIT MODELS OF THE SCHRÖDINGER EQUATIONS, G. Kron. *Physical Review*, New York, N. Y., vol. 67, pt. 1, Jan. 1, 1945, pp. 39–49.


$-f\,\Delta u^1\Delta u^2\Delta u^3\Delta e\,\Delta t$
e
t
u^3
u^2
u^1
$\frac{\Delta e}{2}$
$\frac{\Delta t}{2}$

$$-\frac{\partial}{\partial u^1}a_1\frac{\partial}{\partial u^1}\phi-\frac{\partial}{\partial u^2}a_2\frac{\partial}{\partial u^2}\phi-\frac{\partial}{\partial u^3}a_3\frac{\partial}{\partial u^3}\phi+b\phi+d\frac{\partial}{\partial e}\phi+f\frac{\partial}{\partial t}\phi=c$$

Fig. 11. Five-dimensional circuit with variations along energy and time

2. Numerical Solution of Ordinary and Partial Differential Equations by Means of Equivalent Circuits, G. Kron. *Journal of Applied Physics*, New York, N. Y., vol. 16, pt. 3, March 1945, pp. 172–86.

3. The Elements of Nuclear Reactor Theory, S. Glasstone, M. C. Edlund. D. Van Nostrand Company, Inc., New York, N. Y., 1952.

4. Tensorial Analysis of Integrated Transmission Systems. Part I—The Six Basic Reference Frames, Gabriel Kron. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1239–48.

5. Tensorial Analysis of Integrated Transmission Systems. Part II—Off Nominal Turn Ratios, Gabriel Kron. *AIEE Transactions*, vol. 71, pt. III, 1952, pp. 505–12.

6. Tensorial Analysis of Integrated Transmission Systems. Part III—The Primitive Division, Gabriel Kron. *AIEE Transactions*, vol. 71, pt. III, 1952, pp. 814–21.

7. Tensorial Analysis of Integrated Transmission Systems. Part IV—The Interconnection of Transmission Systems, Gabriel Kron. *AIEE Transactions*, vol. 72, pt. III, 1953, pp. 827–39.

8. A Set of Principles to Interconnect the Solutions of Physical Systems, Gabriel Kron. *Journal of Applied Physics*, New York, N. Y., vol. 24, Aug. 1953, pp. 965–80.

9. A Method to Solve Very Large Physical Systems in Easy Stages, G. Kron. *Proceedings*, Institute of Radio Engineers, New York, N. Y.

10. Equivalent Circuit of the Field Equation of Maxwell—I, G. Kron. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 32, pt. 5, May 1944, pp. 289–99.

11. Equivalent Circuits of Compressible and Incompressible Fluid Flow Fields, G. Kron. *Journal of Aeronautical Sciences*, New York, N. Y., vol. 12, pt. 2, April 1945, pp. 221–35.

12. Equivalent Circuits of the Elastic Field, G. Kron. *Journal of Applied Mechanics*, New York, N. Y., vol. 11, Sept. 1944, pp. 149–61.

◆

No Discussion

Magnetic-Amplifier Circuits with Full-Wave Output and Half-Wave Control Signals

HAROLD W. LORD
FELLOW AIEE

THE advantages of half-wave magnetic-amplifier circuits of the flux reset type have been described in several papers.[1,2] However, when several such amplifiers are cascaded and the output is used to drive a servomotor, the output often is allowed to be a half-wave.[3] This sort of output has proved satisfactory for either d-c motors or 2-phase a-c motors if a small derating of the motor output is accepted.

In some applications it is desirable to obtain from the magnetic amplifier an a-c output having little or no d-c component. In other circuits a complete full-wave rectified direct current is desirable and even may be required.

This paper will show circuits for, and will describe the operating principles of, several magnetic-amplifier circuits which are capable of being controlled by a half-wave type of signal, yet whose outputs are full-wave. The fast speed of response and the simplicity of the half-wave flux reset type of magnetic amplifier therefore may be utilized for the early low-power magnetic-amplifier stages. The half-wave signal is then converted to full-wave in the power-output stage by interconnecting the two inductors of the magnetic amplifier with a "slave" circuit similar to one of the several circuits to be described.

Fundamentals of the Problem and Its Solution

To provide a full-wave output, the output magnetic amplifier must have two or more inductors, consisting of magnetic cores equipped with a-c coils, which are approximately identical both magnetically and electrically. Assuming that both cores are identical and that the a-c coils have equal turns, then a properly controlled single-phase full-wave output will be had if the following condition obtains: the flux density in the one core at the start of the half-cycle during which it exercises control must be the same as that in the other core at the start of the half-cycle during which it in turn exercises control. In the usual flux reset type of circuit, the flux level in a given core can be set only during the half-cycle period in which the rectifier associated with its a-c or gate winding is normally blocking. A half-wave signal of a phase relationship which is suitable for controlling one inductor is therefore unsatisfactory for directly controlling the alternate inductor. The problem therefore may be solved by arranging the circuit so that one inductor, which may be called the "master" inductor, is controlled directly by the half-wave signal. An interconnecting circuit is then provided which will cause the flux in the core of the second, or slave, inductor to be reset during its normal reset half-cycle by the same amount as that set by the signal in the master inductor core. Interconnecting circuits which provide this slave action on the part of the second inductor are preferably unilateral, since the action of the slave inductor should not influence or override the control exercised upon the master inductor by the signal.

Circuits for A-C Output

One of the commonly used magnetic-amplifier circuits for providing an a-c output is the so-called "doubler circuit." Fig. 1 shows the conventional current-controlled doubler circuit for comparison with the half-wave-controlled type of Fig. 2. The latter is essentally the doubler circuit with the control winding omitted on one inductor and the rectifier eliminated from the gate-winding circuit of this same inductor. In spite of these omissions, the half-wave-controlled circuit of Fig. 2 operates in much the same way as the circuit shown in Fig. 1 does, over most of the control range.

In Fig. 2 the 2-winding inductor *A* and rectifier REC_1 form a half-wave self-saturated type of magnetic-amplifier circuit. The control circuit, as shown, has resistor R_C to provide a relatively high impedance to the flow of current in the control circuit caused by voltages induced in the control coil through transformer action from the gate winding. An inductor or a combination of elements such as rectifiers and alternating and/or unidirectional half-wave voltages of proper time phase and polarity could be substituted for R_2 and the control source. Such control circuits are described in references 1 and 2.

Inductor *B* of Fig. 2 is the slave induc-

Paper 54-249, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committe on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 22, 1954; made available for printing April 16, 1954.

Harold W. Lord is with the General Electric Company, Schenectady, N. Y.

Acknowledgment is made of the assistance of P. A. Fessler and E. J. Reagen in setting up the circuits used to demonstrate operations described in this paper and in obtaining the test data and oscillograms.

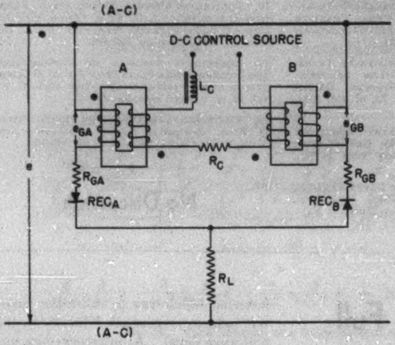


Fig. 1 (left). Conventional doubler circuit

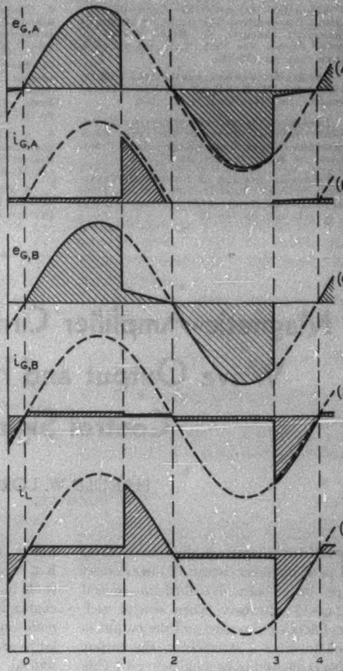


Fig. 3 (right). Idealized wave forms for circuit of Fig. 2

A—Gate voltage of inductor A
B—Gate current of inductor A
C—Gate voltage of inductor B
D—Gate current of inductor B
E—Load current

tor. Its core should be very similar to the core of master inductor *A* both in physical dimensions and in magnetic characteristics. However, it need not match the magnetic characteristics of the core of inductor *A* as closely as is required for current-controlled conventional full-wave self-saturated magnetic-amplifier circuits. Inductor *B* is wound with a gate winding only. This should be wound to the same specifications as the gate winding of inductor *A*. The purpose of the circuit consisting of transformer *T*, resistor R_1, and rectifier REC_1 will be explained later. For the moment, it will be assumed that this circuit is removed by disconnecting at the dashed lines.

The general manner in which inductor *B* acts to provide the desired slave action is quite simple. Note that its winding connects the load resistor R_L directly across the supply voltage. An inductor cannot have a steady-state direct-voltage component across it which is greater than the resistive drop of a direct current flowing through it. This slave inductor therefore must pass such currents as are required to maintain a voltage across the load which is predominantly alternating. Any attempts of the controlled inductor *A* and its associated rectifier REC_A to supply a large direct-voltage component to the load during one half-cycle will be nearly completely offset by a current through inductor *B* during the next half-cycle.

Fig. 3 shows current and voltage wave forms which depict the operation of the circuit of Fig. 2. These wave shapes are derived from idealized rectangular hysteresis loops by a method similar to that described in reference 4. The shaded portions indicate the wave forms when inductor *A* is so controlled as to have a firing angle of 120 degrees. During the prefiring period 0 to 1 when inductor *A* is holding off, the flux in core *B* is being reset by the same amount that was reset in core *A* during the previous half-cycle by the control signal. During this period the gate winding of inductor *B* is for all practical purposes connected in parallel with the gate winding of inductor *A*; hence the volt-time area will be the same for both windings during this period. The only differences during this period are that the magnetic flux in core *A* is changing from the value of reset flux toward saturation and the flux in core *B* is resetting from the residual flux-density value toward a reset value. After inductor *A* fires, the flux in the core of induc-

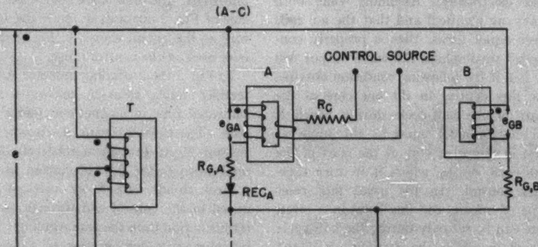


Fig. 2. Half-wave-controlled doubler circuit

tor B continues to reset at a much lower rate, as shown in Fig. 3(C) between points 1 and 2. This rate is determined by the voltage drop across rectifier REC_A and the resistance of the gate winding of inductor A. During the next half-cycle, from points 2 to 3, inductor B will therefore delay firing for a little more than 120 degrees of this half-cycle. During this same time interval the flux in core A can be reset by the signal, as shown in A of Fig. 2. When inductor B fires at point 3 the resetting of the flux in core A is terminated, except for that provided by the resistance voltage drop across inductor B, even though the signal current may continue to try to reset this flux.

If the firing angle of inductor A is advanced or retarded, the firing angle of core B will likewise advance or retard. However, inductor B is not a perfect slave to inductor A. The voltage drop across rectifier REC_A and resistance $R_{G,A}$ causes some resetting of core B even when inductor A fires at 0 degrees, so the firing angle of inductor B will not advance completely to the 180-degree point. Also, when inductor A holds off for a complete half-cycle (cutoff), the exciting current of two cores is flowing through load R_2 during the half-cycle that inductor A holds off. However, rectifier REC_A allows only the exciting current of inductor B to flow during the half-cycle in which inductor B holds off. This requires the firing angle of inductor B, when inductor A is at cutoff, to lag that of inductor A by a little less than 180 degrees in order that the average direct voltage across the load be zero.

Neither of these errors in slave action is large enough to be serious in practical circuits. The delay of firing of inductor B, when inductor A is firing at zero degrees, is small if the forward drop of rectifier REC_A is low. If desired, improved slave action in the cutoff region, and a reduction of approximately two to one in the load current at cutoff, can be had by using the circuit consisting of transformer T, resistor R_1, and rectifier REC_1 shown in Fig. 2. The transformer should have an approximate 1-to-1 turns-ratio between primary and secondary. Resistor R_1 should be of such a resistance value as to pass an average value of current through rectifier REC_1 which approximately equals the average current through rectifier REC_A at cutoff. With the a-c polarities indicated by the large dots in the diagram of Fig. 2, these two rectifiers will be conducting during the same half-cycle of supply voltage. Rectifier REC_1 therefore will reduce the average current in the load during the half-cycle REC_A conducts, thereby reducing the average current that inductor B must pass during the next half-cycle of supply voltage. If resistor R_1 is made too low, the firing of slave inductor B will lag the firing of inductor A more than 180 degrees even in the region of cutoff.

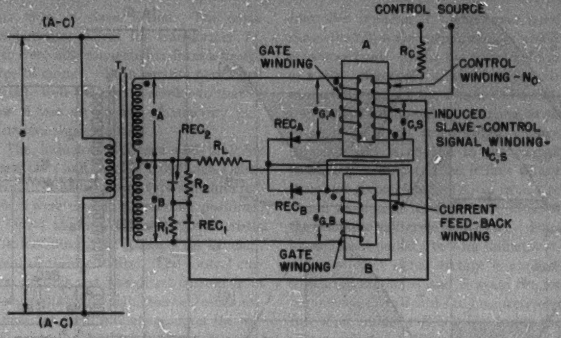


Fig. 4. Full-wave center-tap magnetic amplifier with slave action provided by induced voltage from master inductor

Although the slave action of inductor B responds instantaneously to inductor A except for the required half-cycle of delay time, this circuit has speed-of-response characteristics similar to those exhibited by any a-c output or doubler circuit. Inductor A will respond to a properly phased control signal in the turn-on sense in as short a time as 1/2 cycle. In the turn-off direction, however, the firing of inductor B exerts a loading effect upon the resetting of inductor A. This reduces the speed of response to such an extent that several cycles of power-supply frequency may be required to reach cutoff from the full "on" condition.

Circuits for Full-Wave D-C Output

When the output of a self-saturated magnetic amplifier is direct current, the speed of response is not limited by the loading effects inherent in the a-c output circuits. It is therefore of interest to apply this same slave action to a circuit whose output is normally direct current, rather than to rectify the a-c output of a circuit such as that of Fig. 2.

Fig. 4 shows a full-wave center-tap self-saturated magnetic-amplifier circuit with the flux reset type of slave action applied to it. In this figure inductor A is controlled by any desired means suitable for a half-wave type of circuit, including half-wave signals which occur only during the reset half-cycle for this core. In addition to the customary gate winding and control winding, inductor A includes a winding labeled "induced slave-control winding—$N_{C,S}$." This latter winding has the same number of turns as the gate winding, but the cross-sectional area of conductor needs to be only a few per cent of that of the gate winding. Inductor B is similar to inductor A minus the control windings and may include a current feedback winding of some sort, the purpose of which will be explained later. Rectifiers REC_A and REC_B are the main power-handling rectifiers. Rectifiers REC_1 and REC_2 are units of comparatively low current rating, since they are in the control circuit. Their functions, along with those of resistors R_1 and R_2, will be explained later.

Fig. 5 indicates the voltage and current wave forms of certain parts of the circuit when inductor A is controlled during its reset half-cycle so as to fire at a phase angle of approximately 120 degrees.

Fig. 5(A) shows that during the period 0 to 1, when inductor A is holding off, there will be a voltage induced in winding $N_{C,S}$ whose volt-time area equals that of the gate winding during the hold-off period. This, in turn, is shown to be equal to the volt-time area of the voltage induced in these winding during the reset half-cycle of this inductor. One end of winding $N_{C,S}$ is connected directly to one end of the gate winding of inductor B, and the other end of winding $N_{C,S}$ connects through rectifier REC_1 and resistor R_1 to the other end of the gate winding of inductor B. Rectifier REC_1 is so polarized as to permit the voltage of winding $N_{C,S}$ during the hold-off period 0 to 1 of inductor A, to impress a resetting voltage across the gate winding of inductor B. With both of the gate windings and winding $N_{C,S}$ having equal turns, and neglect-

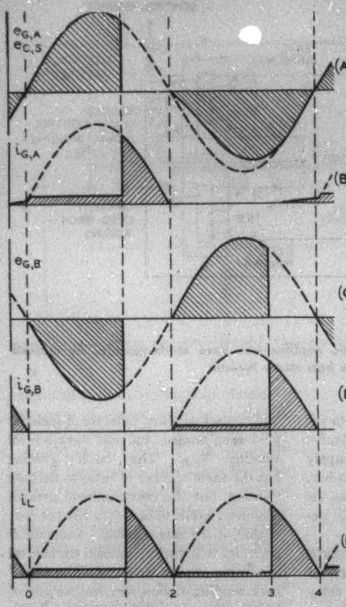


Fig. 5 (left). Idealized wave forms for circuit of Fig. 2

A—Gate and induced control voltages of inductor A
B—Gate current of inductor A
C—Gate voltage of inductor B
D—Gate current of inductor B
E—Load current

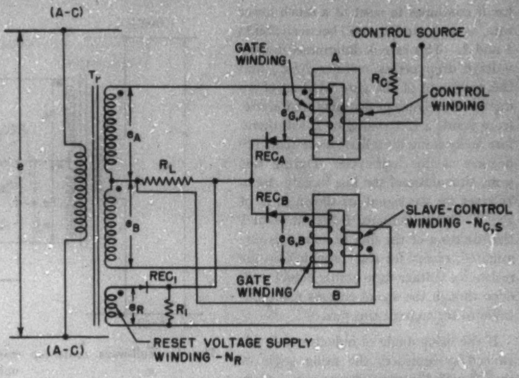


Fig. 6. Full-wave center-tap magnetic amplifier with slave action controlled by load voltage

ing for the moment the effect of resistance drops, the volt-time area of reset supplied to inductor *B* during its reset period just equals the hold-off volt-time area of inductor *A* as shown in Fig. 5(C), points 0 to 1. When it becomes inductor *B*'s turn to control the current to the load during the next half-cycle of supply voltage (rectifier REC_B conducting), inductor *B* will delay its firing, so as to fire 180 degrees after the firing angle of inductor *A*, thus providing the desired slave action.

Rectifier REC_1 prevents loading of winding $N_{C,S}$ during the reset period of inductor *A*, points 2 to 3, but it will not prevent the hold-off voltage of the inductor *B* gate winding from causing an unwanted reset of inductor *A*. Such a "kickback" action is successfully prevented by conduction of rectifier REC_2 during this half-cycle, points 2 to 4. This applies an inverse voltage to rectifier REC_1 approximately equal to e_B, in addition to that from winding $N_{C,S}$. Then, since voltage $e_{G,B}$ cannot exceed voltage e_B, rectifier REC_1 will not conduct during this half-cycle even though the voltage of winding $N_{C,S}$ during this period is zero. The purpose of resistor R_1 is to limit the current through rectifier REC_2 so that the voltage across rectifier REC_2 in the conducting direction will be small compared with that across resistor R_1.

Although voltage drops during the resetting of inductor *B* were neglected in this explanation, their effects cannot be ignored. The flow of inductor *B* reset exciting current through the forward resistance of rectifier REC_1, resistor R_1, and the winding resistances causes a voltage drop which reduces the amount that inductor *B* will be reset by the resetting induced voltage in winding $N_{C,S}$ when inductor *A* is holding off. The effect of this voltage drop is to cause the firing angle of inductor *B* to lag that of inductor *A* by less than 180 degrees. Resistor R_2 provides a simple means for compensating for such effects. If it is of the proper value, then a small sinusoidal component of current will be fed in to resistor R_1 of the correct polarity to compensate for the resistive voltage drops caused by inductor *B* reset exciting current. A proper value for resistor R_2 is indicated when the inductor *B* firing angle lags that of inductor *A* by 180 degrees as cutoff.

It is well known to those familiar with self-saturated magnetic amplifiers that rectifier leakage currents tend to reset the cores of their associated inductors and thereby to act as a negative feedback. In this circuit their effect is to cause the firing angle of inductor *B* to lag that of inductor *A* by more than 180 degrees. Thus, when inductor *A* reaches maximum output (firing angle of 0 degrees), inductor *B* will not have reached full output. This effect may be corrected by passing the load current through a series-connected feedback winding of sufficient turns to provide a positive feedback which equals the negative feedback caused by rectifier leakage. Similar compensation may be had by a shunt feedback circuit when the application involves a variable load impedance. A feedback winding, having turns that fall within the range of 5 to 10 per cent of the turns of the gate winding, when connected in series with a suitable resistor and this series circuit connected across the load, will provide satisfactory performance without excessive loading effects because of induced voltages in the feedback winding.

Slave action of the second inductor may be had by a somewhat different mode of operation, as shown in Fig. 6. An auxiliary secondary winding on the power transformer T_r, identified as "reset voltage supply winding—N_R," supplies a resetting voltage to "slave-control winding—$N_{C,S}$" on inductor *B* during the normal reset half-cycle for inductor *B*. With the a-c polarities as indicated by large dots, the current required to reset inductor *B* flows in the series circuit composed of the forward or conduction direction through the small control rectifier REC_1, the load as represented by R_L, and the winding $N_{C,S}$.

During the prefiring interval of the conduction period for REC_A, period 0 to 1 in Fig. 5(A), the voltage across load R_L is relatively low. A large portion of voltage e_R will then be impressed across winding $N_{C,S}$ to reset inductor *B*.

After inductor *A* fires, period 1 to 2

on Fig. 5, the voltage across load R_L acts to reduce the net voltage across winding $N_{C.S}$. If the voltage e_R is made such that, after inductor *A* fires, the instantaneous voltage across load R_L always exceeds that of e_R, an inverse voltage will appear across rectifier REC_1 during the period 1 to 2. As a result, no further resetting of inductor *B* will occur for the remainder of the half-cycle.

By a proper choice of the turns of winding $N_{C.S}$ the flux in core *B* can be made to reset during the period 0 to 1 by an amount which will delay the firing angle of inductor *B* so as to lag the firing angle of inductor *A* by 180 degrees. The turns of winding $N_{C.S}$ will be approximately correct if the turns-ratio of the gate winding to winding $N_{C.S}$ is 110 per cent of the ratio of e_B to e_R. It is better to have this ratio a little too high than too low. Any tedency of inductor *B* to reset too rapidly will cause conduction of rectifier REC_B during the reset period 0 to 1 in Fig. 5. This will cause the current in the reset circuit to increase, and the voltage supplied by winding N_R, which is in excess of that required to reset inductor *B* properly, will appear as additional voltage drops across series-resistive elements of the slave control circuit.

Leakage currents of rectifier REC_B will impair the operation of this circuit in the same way as that described for the circuit of Fig. 4. Resistor R_1 in the control circuit of Fig. 6, in conjunction with winding $N_{C.S}$, provides a shunt-type positive feedback circuit for inductor *B* for opposing the negative feedback effect of the leakage current of rectifier REC_B.

Both of the modes of slave operation which have been described for the full-wave center-tap type of circuit can be used equally well for the full-wave bridge type of self-saturated magnetic-amplifier circuit.

Experimental Results

A circuit similar to that of Fig. 4 was set up to demonstrate the operation of this type of slave action. The same components, except for transformer T_r, also were interconnected to form a bridge type of circuit. Both circuits performed equally well. A circuit similar to that of Fig. 6 was set up and satisfactory performance was obtained.

The inductors designed and built for these tests had cores of 0.002 inch Deltamax. Standard power-type selenium rectifiers were used in all rectifier positions.

Fig. 7 shows a transfer characteristic obtained during the course of the development of these circuits. The control current for this test was direct current from a source having a high internal resistance. Effects of rectifier leakage upon the signal-controlled inductor are here apparent, causing a shearing of the top of the curve to the right[3] to such an extent that positive signal current is required over much of the control range.

Since half-wave control and speed of response were the main considerations in this development, an electronic control circuit was used to provide a rapidly varying signal to control this magnetic amplifier. This control circuit provided complete half-cycles and/or portions of half-cycles, but only during the normal reset half-cycle of the signal-controlled inductor. It included a means for varying the signal rapidly between limits of full "on" and cutoff and at accurately timed repetition rates which were readily synchronized with respect to the 60-cycle a-c supply. A half-wave positive-bias circuit assured full conduction of the master inductor in the absence of a resetting signal.

The results of a speed-of-response test of one of the half-wave-controlled full-wave d-c output magnetic-amplifier circuits are shown in Fig. 8. This oscillogram was obtained by a dual-beam cathode-ray oscilloscope with synchronized sweeps. The upper trace shows the half-wave resetting signal applied to signal-controlled inductor *A*. The lower trace shows the full-wave output voltage with a resistance load. Using the signal voltage as a time reference, with the time periods as shown by letters, the sequence of operation may be followed by starting at *a*. From *a* to *b* no reset signal voltage (upper trace) is applied to the control winding of inductor *A*, so it fires (lower trace) for a complete half-cycle from *b* to *c*, and slave inductor *B* likewise fires for the complete half-cycle from *c* to *d*. During the reset period *c* to *d* for inductor *A*, again no reset signal voltage is applied, so both inductors *A* and *B* again fire for the intervals *c* to *d* and *e* to *f* respectively. But while inductor *B* is firing during the period *e* to *f*, a half-cycle of reset signal voltage is applied to the control winding as shown by the upper trace. Consequently, master inductor *A* holds off for practically the whole half-cycle for period *f* to *g* (lower trace), and slave inductor *B* duplicates this performance from *g* to *h*. The upper trace shows that inductor *A* again is reset during period *g* to *h*, so again both inductors hold off for the complete half-cycles of *h* to *i* and *i* to *j* (lower trace). However, during period *i* to *j*, the upper trace shows that master inductor *A* received a resetting control voltage for approximately 1/4 cycle, so it is reset only partially. As a result, the lower trace shows that master inductor *A* fires for the last half of the half-cycle period from *j* to *a*, and slave inductor *B* does nearly the same during period *a* to *b*. This same sequence of events was repeated in this test for every 5 cycles of supply frequency, corresponding to a repetition rate of 12 per second.

The preceding explanation of the sequence of events shown in Fig. 8 clearly demonstrates that, in this full-wave output circuit, the master inductor responds to a signal with the half-cycle-per-stage time delay characteristic of the flux reset

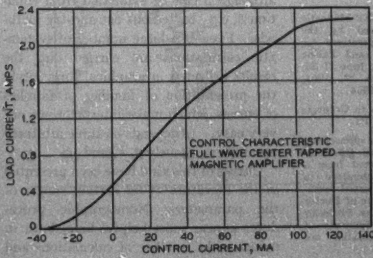


Fig. 7 (left). Transfer characteristic of circuit similar to Fig. 4

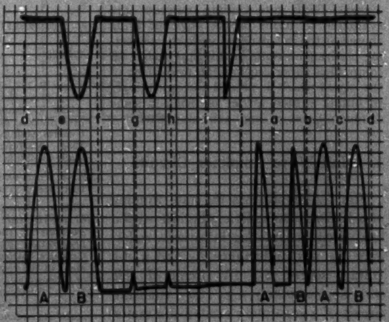


Fig. 8 (right). Oscillogram showing speed of response provided by circuit of Fig. 4

type of magnetic amplifiers. The slave inductor is shown to require only an additional half-cycle of time delay for completing its slave action. The oscillogram also shows the response time to be the same for both increasing and decreasing outputs.

References

1. On the Mechanics of Magnetic Amplifier Operation, Robert A. Ramey. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1214–23.

2. On the Control of Magnetic Amplifiers, R. A. Ramey. *Ibid.*, pp. 2124–28.

3. An Improved Magnetic Servo Amplifier, C. W. Lufcy, A. E. Schmid, P. W. Barnhart. *AIEE Transactions*, vol. 71, pt. I, Sept. 1952 pp. 281–89.

4. The Influence of Magnetic Amplifier Circuitry Upon the Operating Hysteresis Loops, Harold W. Lord. *AIEE Transactions*, vol. 72, pt. I, 1953 (Jan. 1954 section), pp. 721–28.

5. Theory of Magnetic Amplifiers with Square-Loop Core Materials, H. F. Storm. *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 629–40.

No Discussion

The Operation of Magnetic Amplifiers with Various Types of Load

Part I—Load Currents for Given Angle of Firing

L. A. FINZI
MEMBER AIEE

R. R. JACKSON
ASSOCIATE MEMBER AIEE

Synopsis: The operation of magnetic amplifiers with sensitive core materials is characterized by time intervals in the cycle during which the voltage of the a-c power source is balanced (or "absorbed") almost entirely by the rates of change of core fluxes, and by other intervals in which the core fluxes are nearly constant and the voltage of the power source substantially appears applied to the load.

Transitions from the first to the second mode are very abrupt. Thus load currents can be evaluated as if the voltage of the a-c power source were periodically applied at some time or "angle of firing" by the firing of a thyratron or by the closing of a switch and removed at some later time. A very simple solution is obtained if the output current upon firing is limited by nothing else but resistances. The analysis, however, becomes much more elaborate if this restriction is removed.

This paper examines instances of practical interest in which the load circuit includes resistances, inductances and also d-c voltage sources, in conjunction with rectifying elements variously inserted. Load currents and gate-winding currents are evaluated in terms of the angle of firing chosen as the independent variable, neglecting prefiring currents in this approach. (A companion paper aims at the evaluation of the signal which is needed correspondingly to obtain wanted angles of firing in some common types of high-gain amplifiers.[1])

1. Half-Wave Amplifier Circuits

Fig. 1(A) represents the gate winding of a half-wave magnetic amplifier with a load circuit containing in series a resistance R_L, and inductance L_L and a d-c voltage source E_L. (A circuit of this sort approximates, e.g., the armature circuit of a separately excited d-c motor at constant speed.) R_g is the resistance of the amplifier gate winding and gate rectifier and L_g is its saturated inductance. Thus the total resistance and inductance seen by the a-c voltage source $v_g = V_{gm} \sin \omega t$ upon firing are $R = R_L + R_g$ and $L = L_L + L_g$. (If the rectifier has appreciable incremental resistance and threshold voltage in the forward direction, these effects can be accounted for by increasing R_g and modifying E_L.)

With $k = E_L/V_{gm}$, where $-1 \leq k \leq 1$, the differential equation of the circuit in the saturated mode is simply

$$V_{gm}(\sin \omega t - k) = Ri_L + L\frac{di_L}{dt} \qquad (1)$$

Thus for firing at $\omega t = \alpha$, under the assumption of negligible prefiring magnetizing current

$$i_L = \frac{V_{gm}}{R}\{[\cos \theta \sin(\omega t - \theta) - k] - [\cos \theta \sin(\alpha - \theta) - k]\epsilon^{-(\omega t - \alpha)\cot \theta}\} \qquad (2)$$

where

$$\tan \theta = \frac{\omega L}{R}$$

This expression is valid over the "saturation" interval initiated at α and terminated at the "extinction angle" β. At $\omega t = \beta$, i_L becomes zero. (The gate rectifier excludes the possibility of negative values for i_L.) Therefore the angle β is defined by the transcendental equation

$$0 = [\cos \theta \sin(\beta - \theta) - k] - [\cos \theta \sin(\alpha - \theta) - k]\epsilon^{-(\beta - \alpha)\cot \theta} \qquad (3)$$

The cyclic average of output current is found as

$$I_{Ls} = \frac{1}{2\pi}\int_{\alpha}^{\beta} i_L d\omega t = \frac{V_{gm}}{2\pi R}[(\cos \alpha - \cos \beta) - k(\beta - \alpha)] \qquad (4)$$

where the additional subscript s is introduced to remind that contributions of prefiring currents have been neglected.

The form of equation 4 appears well justified on the basis of volt-time area concepts. During the saturation interval the magnetic core does not absorb any voltage, and the voltage of the a-c power supply is absorbed partly by the battery E_L and partly by the passive elements R and L. But the cyclic volt-radian area absorbed by L is zero (as $i_L = 0$ at $\omega t = \alpha$ and again at $\omega t = \beta$). Hence the volt-radian area $2\pi RI_{Ls}$ absorbed by R results simply as the volt-radian area $V_{gm}(\cos \alpha - \cos \beta)$ of the power source minus the volt-radian area $E_L(\beta - \alpha)$ absorbed by the battery.

An explicit solution for β is not possible, although β can be evaluated from equation 3, e.g., by tedious cut-and-try methods. Even if a large number of systematic calculations are carried out, the results collected in graphical form imply the presentation of families of families of curves, which make interpolations not very easy. (β depends on three arbitrary parameters, namely θ, k, and α.)

Curves of this kind have been presented in the literature[2-4] for certain ranges of the parameters. Nomographic procedures appear most helpful here, both in reducing the number of calculations and

---

Paper 54-237, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 23, 1954; made available for printing April 15, 1954.

L. A. Finzi and R. R. Jackson are with Carnegie Institute of Technology, Pittsburgh, Pa.

This work was performed in part under sponsorship of the United States Office of Naval Research. Portions of the paper were abstracted from a dissertation by R. R. Jackson, submitted to Carnegie Institute of Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy. Helpful suggestions by Professor H. J. Greenberg of the Carnegie Institute of Technology Department of Mathematics are gratefully acknowledged.

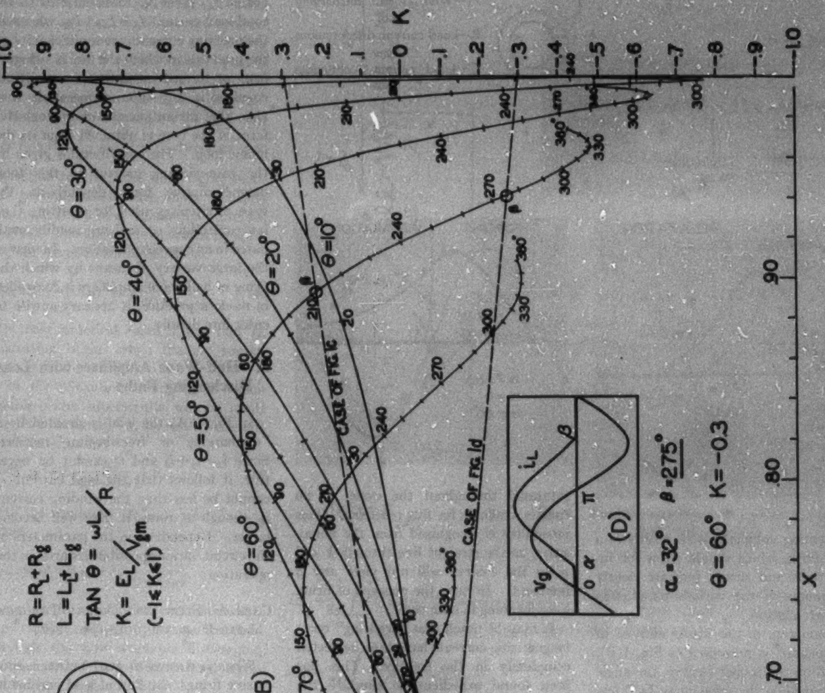
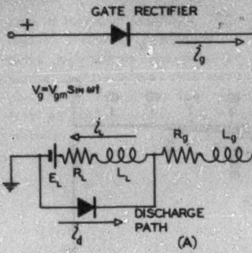
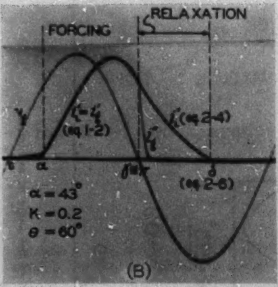


Fig. 1. Half-wave amplifier circuit. Determination of the angle of extinction





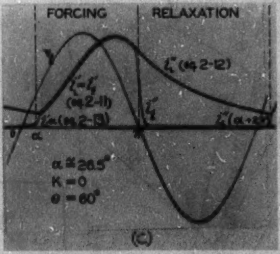


Fig. 2. The half-wave amplifier

A—With load discharging path
B—Load current, discontinuous flow
C—Load current, continuous flow

in presenting comprehensive results in a unified form, which should allow for interpolations and should indicate clearly how changes of the various parameters affect the solution.

A nomogram of this kind,[5] derived in the Appendix, is presented in Fig. 1(B). Given α on the proper θ-curve, the angle β is obtained on the same curve by interception with a straight line joining the given α-point with the assigned k-point on the k-axis. Two examples are shown in dotted lines for $\theta = 60$ degrees and $k = +0.3$ and $k = -0.3$; corresponding wave forms of load currents are seen in the oscillograms of Fig. 1(C) and (D). (The meaning of the x- and y-axes shown is discussed in the Appendix with additional comments. No use is made of these axes in the evaluation of β.)

With the polarities and sign conventions chosen in Fig. 1(A), k is positive when E_L opposes the flow of positive current. Situations may occur in which E_L has opposite polarity, i.e., aiding the flow of current. The solution given by the nomogram and by equation 4 remains quite general, with the consideration that E_L and thus k are negative numbers in such cases. Instantaneous values of i_L are shown in Fig. 1(D) for situations of this kind.

With k sufficiently negative and L large enough, it is possible that no extinction occurs. The amplifier core then remains saturated throughout the cycle, as no time is available for flux resetting. This possibility is recognized from the nomogram, as the straight line through k and α on the θ-curve will not yield any β intercept. In fact, the concept of firing loses meaning in such cases.

It may be noted that "prefiring" core-magnetizing currents have been neglected completely in the solution. This has been found expedient to simplify the integration of equation 1. Corrections can be introducted evaluating the contribution I_{LP} given by these currents to the total load current $I_L = I_{LS} + I_{LP}$ whenever the prefiring currents are significant; e.g., the intervals in which the flux is swinging from or toward saturation are easily recognized in any type of operation, thus also the instantaneous magnetomotive force of the core is obtained from its dynamic loop. The contributions given by the gate-winding current to this total magnetomotive force depend on the type of circuitry used for resetting (i.e., for controlling α) and are readily evaluated in any specific situation. In view of the large variety of means by which the firing of half-wave amplifiers is controlled in modern practice, it appears unwise to enter into details.

2. Half-Wave Amplifiers with Load Discharging Paths

In Fig. 2(A) the load is shunted by a "discharging" or "freewheeling" rectifier. Since $i_L = i_g + i_d$ and i_d cannot be negative, it follows that the load current i_L cannot be less than the winding current i_g, though at times it may well become larger. Depending on the parameters of the circuit various possibilities arise in the operation:

CASE A—FLOW OF CURRENT THROUGH LOAD DISCONTINUOUS, FIG. 2(B)

When i_L decays to zero between successive firings, the flow of load current is

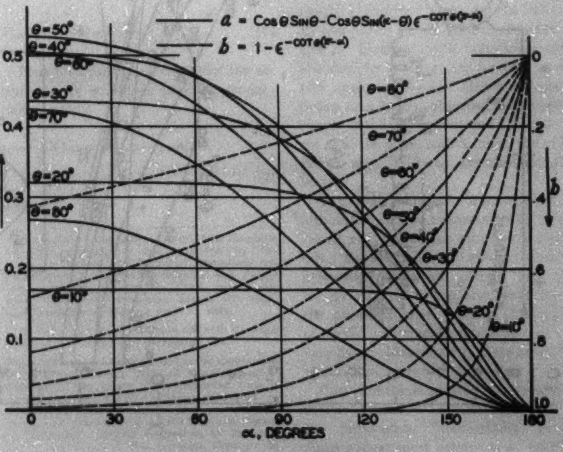


Fig. 3. Quantities a and b of equation 9

called discontinuous. Since at $\omega t=\alpha$ the load current is zero, equation 2 describes the initial transient upon firing. The discharging rectifier does not yet conduct and the load current equals the winding current. This is a first "forcing" part of the transient. (A prime superscript will be used for currents during this interval.)

However, at some angle γ the load voltage $v_L=R_Li_L'+L_L\frac{di_L'}{dt}+E_L$ becomes zero trying to reverse; the discharging rectifier becomes conducting, short-circuiting the load, and a new "relaxation" interval begins. The angle γ is obtained by solving the transcendental equation

$$V_{gm}\sin\gamma=\left(R_gi_L'+L_g\frac{di_L'}{dt}\right)_{t=\gamma/\omega} \tag{5}$$

(In most practical cases $\gamma\cong\pi$, that is, relaxation begins when the a-c supply voltage is about zero.)

In the relaxation interval (for which double-prime superscripts will be used) the load current i_L'' is defined by the differential equation

$$0=R_Li_L''+L_L\frac{di_L''}{d(t-\gamma/\omega)}+E_L \tag{6}$$

with the condition $(i_L''{}_{\omega t=\gamma}=)(i_L')_{\omega t=\gamma}$, while

$$V_{gm}\sin\omega t=R_gi_g''+L_g\frac{di_g''}{d(t-\gamma/\omega)} \tag{7}$$

describes the rapidly decaying gate winding current, which by-passes the load through the now available discharging rectifier path. Its further reversal is prevented by the gate rectifier.

Integration of equation 6 yields

$$i_L''=\left(i_{L\gamma}'+\frac{E_L}{R_L}\right)\epsilon^{-\frac{R_L}{\omega L_L}(\omega t-\gamma)}-\frac{E_L}{R_L} \tag{8}$$

At some angle $\omega t=\delta$, i_L'' becomes zero, and relaxation terminates.

The approximation $\gamma\cong\pi$ allows ready calculation of $i_{L\gamma}'$ and δ, namely

$$i_{L\gamma}'\cong=i_{L\pi}'=\frac{V_{gm}}{R}[\cos\theta\sin\theta-\cos\theta\sin(\alpha-\theta)\epsilon^{-(\pi-\alpha)\cot\theta}-k(1-\epsilon^{-(\pi-\alpha)\cot\theta})]$$
$$=\frac{V_{gm}}{R}(a-kb) \tag{9}$$

where a and b are both functions of θ and α, plotted for the sake of convenience in the family of curves of Fig. 3, and

$$\delta\cong\frac{\omega L_L}{R_L}\ln\left(\frac{R_Li_{L\pi}'}{E_L}+1\right)+\pi \tag{10}$$

The cyclic average current is obtained as the sum

$$I_{LS}=I_{LS}'+I_{LS}'' \tag{11}$$

where

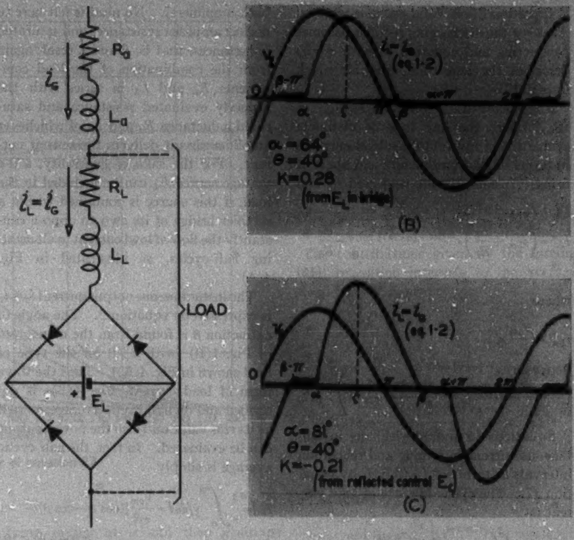


Fig. 4. Two-core amplifiers with a-c output; nonrectified load inductance

$$I_{LS}'=\frac{1}{2\pi}\int_\alpha^\pi i_L'd\omega t=[V_{gm}(\cos\alpha-\cos\pi)-E_L(\pi-\alpha)-\omega Li_{L\pi}']/2\pi R \tag{12}$$

$$I_{LS}''=\frac{1}{2\pi}\int_\pi^\delta i_L''d\omega t=[\omega L_Li_{L\pi}'-E_L(\delta-\pi)]/2\pi R \tag{13}$$

These expressions indicate that the volt-radian area of the a-c source, not absorbed by the core during the forcing interval from α to π, equals the volt-radian area $2\pi RI_{LS}'$ plus the volt-radian area $E_L(\pi-\alpha)$ absorbed by the battery, plus volt-radian area $\omega Li_{L\pi}'$ absorbed by the circuit inductances. In the second part of the transient the volt-radian area $\omega L_Li_{L\pi}'$ of the load inductance is converted into the volt-radian area $2\pi R_LI_{LS}''$ plus the volt-radian area $E(\delta-\pi)$ further absorbed by the battery over the interval π to δ.

If the resistance and saturated inductance of the gate winding are neglected equation 11 simplifies into

$$I_{LS}=\frac{V_{gm}}{2\pi R}[(1+\cos\alpha)-k(\delta-\alpha)] \tag{14}$$

It may be noted that for certain angles of firing, relaxation may not take place at all during the cycle if E_L is positive and sufficiently large. The occurence of this subcase may be recognized from the nomogram of Fig. 1(B) whenever this yields $\beta<\pi$. Then equations 2 and 4 apply throughout.

CASE B—FLOW OF CURRENT THROUGH LOAD CONTINUOUS, FIG. 2(C)

In many instances (e.g., zero or negative values of E_L or large inductances, or early firing in the cycle) equation 10 of the preceding analysis yields $\delta>\alpha+2\pi$, that is, a new firing would seem to occur before the load current of the preceding relaxation interval has died out, in contradiction to the initial condition of zero load current used in the integration of equation 1. In this case the differential equation 1 describes the forcing part of the transient but the solution is modified by the existence of an initial current $i_{L\alpha}'>0$.

To obtain workable results it appears expedient to neglect the resistance R_g and the saturated inductance L_g of the amplifier winding. As a first consequence of this approximation i_g reaches immediately the value $i_{L\alpha}'$ upon core saturation. Thus equation 2 modifies into the following

$$i_L'=\frac{V_{gm}}{R_L}\{[\cos\theta\sin(\omega t-\theta)-k]-[\cos\theta\sin(\alpha-\theta)-k]\epsilon^{-(\omega t-\alpha)\cot\theta}\}+i_{L\alpha}'\epsilon^{-(\omega t-\alpha)\cot\theta} \tag{15}$$

Equation 15 is valid from α until π. At $\omega t=\pi$ the discharging rectifier becomes conducting and equation 8 (with $\gamma=\pi$) describes the relaxation interval, which lasts until $\omega t=\alpha+2\pi$.

The steady-state condition $i_{L(\alpha+2\pi)}''=i_{L(\alpha)}'$ offers the key for the complete evaluation of instantaneous load currents in terms of load parameters and angle of firing. In fact, having neglected R_g and L_g

$$i_L''=\left(i_{L\alpha}'+\frac{E_L}{R_L}\right)\epsilon^{-[\omega t-(\alpha+2\pi)]\cot\theta}-\frac{E_L}{R_L} \tag{16}$$

where

$$i_{L\alpha}'=\frac{V_{gm}}{R_L}\times \left(\frac{\cos\theta\sin\theta\,\epsilon^{-(\alpha+\pi)\cot\theta}-\cos\theta\sin(\alpha-\theta)\epsilon^{-2\pi\cot\theta}}{1-\epsilon^{-2\pi\cot\theta}}-k\right) \tag{17}$$

Equations 15 and 16 yield the instantaneous current in forcing and relaxation intervals in terms of $i_{L\alpha}'$ given by equation 17. The cyclic average of load current is

$$I_{LS}=\frac{1}{2\pi}\int_\alpha^\pi i_L'd\omega t+\frac{1}{2\pi}\int_\pi^{\alpha+2\pi} i_L''d\omega t$$
$$=I_L'+I_L''=\frac{V_{gm}}{2\pi R_L}[(1+\cos\alpha)-2\pi k] \tag{18}$$

This expression is obviously obtainable simply from volt-time area considerations. Equation 18 together with the recognition that the core is saturated only from α to π, supplies sufficient information for relating cyclic average output to control in most common half-wave amplifier circuits. On the other hand, equations 15 to 17 are needed if the wave form of load current is of interest.

3. Two-Core Amplifiers with A-C Output

The analysis of section 1 is extended readily to resistive-inductive loads supplied directly by 2-core amplifiers with a-c output (e.g., doubler or external feedback amplifiers). No need is felt here to neglect amplifier resistances and saturated inductances, that is, R and L result again from the combination of the load components R_L and L_L in series with the properly evaluated resistance and saturated inductance R_a and L_a of whichever amplifier circuit delivers the output current. For the sake of generality, a d-c voltage source E_L can be included in the load, if this source is contained within a rectifier bridge of its own to oppose constantly the flow of load current in alternating half-cycles, as illustrated in Fig. 4(A).

The instantaneous output current $i_G=i_L$ is expressed by equation 2. The angle of extinction β is found from the nomogram of Fig. 1(B) (with $k\geqslant 0$ for the type of load shown in Fig. 4(A)). Thus the wave form of load current, Fig. 4(B), is described and its half-cyclic average, as well as its rms value and also the power output can be evaluated. In fact, the half-cyclic average is simply

$$I_{LS}=\frac{1}{\pi}\int_\alpha^\beta i_L d\omega t=\frac{V_{gm}}{\pi R}[(\cos\alpha-\cos\beta)-k(\beta-\alpha)] \tag{19}$$

Since firing occurs at α and then again at $\alpha+\pi$ it is evident that α cannot be advanced in the half-cycle below a value α^* for which $\beta=\alpha^*+\pi$. In this limiting condition (which can be easily determined for given k and θ on the nomogram by pivoting a straight edge about k) the amplifier does not absorb the a-c supply voltage at any time. The half-cyclic load current is maximum and is

$$I_{L\,\max}=\frac{V_{gm}}{\pi R}(2\cos\alpha^*-k\pi) \tag{20}$$

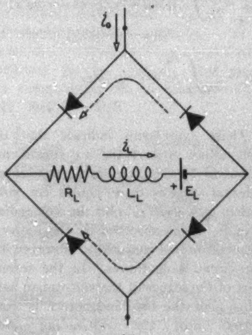
If $k=0$ equation 3 yields $\alpha^*=\theta$ and

$$I_{L\,\max}=\frac{2}{\pi}\,\frac{V_{gm}}{\sqrt{R^2+\omega^2L^2}}$$

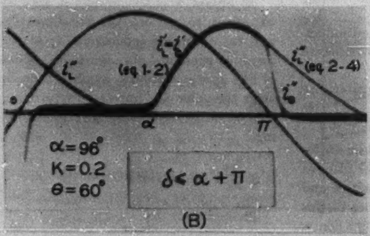
On the other hand, the analysis of section 1 can be extended to some 2-core situations in which it may be said that a negative k appears in the output circuit. This occurs in simple saturable reactors (e.g., with series-connected gate windings, as obtained from Fig. 8(A) under elimination of the feedback windings) as the control voltage E_c as well as the total control circuit resistance R_c and inductance L_c appear reflected into the gate circuit by the current transformer action of the core, which is unsaturated. With the initial sign conventions, k here is negative as E_c is reflected with polarity aiding the flow of current in either saturation interval in the cycle, namely, β is found from the nomogram of Fig. 1(B) with

$$k=-\frac{|E_c|}{V_{gm}}\,\frac{N_g}{N_c}$$

R and L are the resistances and inductances of the output circuit modified by addition of the reflected resistances and inductances of the control circuit. More precisely, R_a is then the resistance of both gate windings plus $R_cN_g^2/N_c^2$; L_a is four times the saturated inductance of one gate winding plus conceivable additional inductances of the control circuit multiplied times N_g^2/N_c^2. Thus the wave forms of currents, as obtained by others,[6] are defined completely, Fig. 4(C) and the influence of the various parameters of gate and control is evidenced. In particular, equations 19 and 20 still apply.

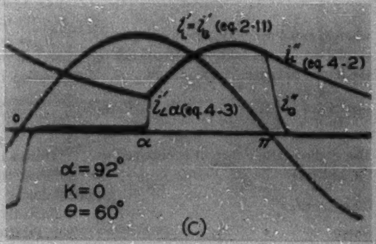


(A)



(B)

Fig. 5. Two-core amplifiers. Rectified load



(C)

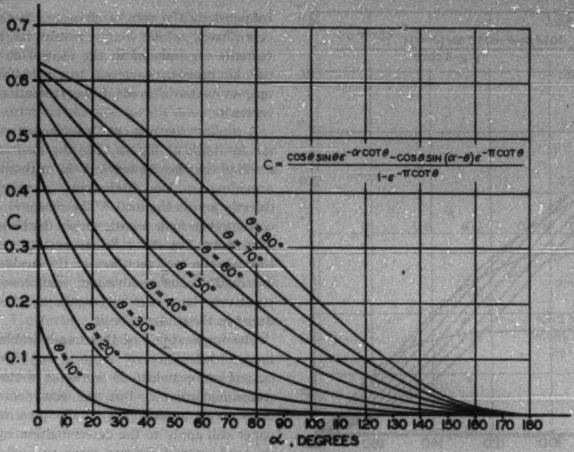


Fig. 6. Quantity c of equation 23

Moderate extensions of this treatment can be made if feedback turns are added to the saturable reactor, as shown in Fig. 8(A), under consideration that the transformer ratio applied in the reflection of control circuit quantities becomes $(N_g-N_f)/N_c$. Phenomena of commutation within the feedback loop limit this extension to cases of comparatively low control circuit impedances.

In Figs. 4(B) and (C) the load current reaches its peak value at ζ. Knowledge of this angle is of interest in some situations; as shown in the companion paper,[1] this is the case, for instance, in the evaluation of control current requirements for certain types of 2-core amplifiers. Graphical means for a quick determination of ζ are given in the Appendix.

4. Two-Core Amplifiers with A-C Output and Externally Rectified Load

Unidirectional load currents are obtained from magnetic amplifiers with a-c output by enclosing the whole load (resistive-inductive, with d-c voltage source) within a rectifier bridge, Fig. 5(A). In this case the instantaneous load current cannot be less than the absolute value of the amplifier output current, though at times it may well become larger, as discharging paths are offered by the four arms of the bridge in simultaneous conduction. The analyses in section 2 are modified here under the recognition that firing occurs at α and then again at $\alpha+\pi$.

The analysis of case *A*, of discontinuous flow applies as long as equation 10 yields load current extinction at $\delta\leq\alpha+\pi$. The instantaneous currents i_L' and i_L'' during forcing and relaxation are still expressed by equations 2 and 8, and shown in Fig. 5(B). Also, the current $i_{L\pi}'$ at the instant of transition from forcing to relaxation is given by equation 9.

But, as two firings take place now in one cycle of the power supply, the average load I_L is evaluated now as $I_{LS}=I_L'+I_L''$ where I_L' and I_L'' are twice the values given by equations 12 and 13. If resistances R_a and inductances L_a of gate windings (accounting also for reflected quantities from other coupled amplifier circuits) are negligible

$$I_{LS}=\frac{V_{gm}}{\pi R_L}[(1+\cos\alpha)-k(\delta-\alpha)] \qquad (21)$$

On the other hand, if equation 10 yields $\delta>\alpha+\pi$ the analysis of case *B* of continuous flow applies, with the expedient approximation $R_a=L_a=0$. The matching condition of equal load currents at two successive firings is stated here as $i_{L(\alpha+\pi)}''=i_{L\alpha}'$.

Thus the instantaneous load currents during forcing and relaxation, as shown in Fig. 5(C) are

$$i_L'=\frac{V_{gm}}{R_L}\{[\cos\theta\sin(\omega t-\theta)-k]-[\cos\theta\sin(\alpha-\theta)-k]\epsilon^{-(\omega t-\alpha)\cot\theta}\}+i_{L\alpha}'\epsilon^{-(\omega t-\alpha)\cot\theta} \qquad (15)$$

$$i_L''=\left(i_{L\alpha}'+\frac{E_L}{R_L}\right)\epsilon^{-[\omega t-(\alpha+\pi)]\cot\theta}-\frac{E_L}{R_L} \qquad (22)$$

where

$$i_{L\alpha}'=\frac{V_{gm}}{R_L}\times\left(\frac{\cos\theta\sin\theta\epsilon^{-\alpha\cot\theta}-\cos\theta\sin(\alpha-\theta)\epsilon^{-\pi\cot\theta}}{1-\epsilon^{-\pi\cot\theta}}-k\right)=\frac{V_{gm}}{R_L}(c-k) \qquad (23)$$

The contributions given by the forcing and relaxation intervals to the average load current I_L are

$$I_{LS}'=\frac{1}{\pi}\int_\alpha^\pi i_L'd\omega t=\frac{V_{gm}(1+\cos\alpha)-\omega L\left(i_{L\alpha}'+\frac{E_L}{R_L}\right)\times(\epsilon^{\alpha\cot\theta}-1)-E_L(\pi-\alpha)}{\pi R_L} \qquad (24)$$

$$I_{LS}''=\frac{1}{\pi}\int_\pi^{\alpha+\pi} i_L''d\omega t=\frac{\omega L\left(i_{L\alpha}'+\frac{E_L}{R_L}\right)(\epsilon^{\alpha\cot\theta}-1)-E_L\alpha}{\pi R_L} \qquad (25)$$

and thus

$$I_{LS}=I_{LS}'+I_{LS}''=\frac{V_{gm}}{\pi R_L}[(1+\cos\alpha)-k\pi] \qquad (26)$$

where k is a negative number if E_L aids the flow of load current.

Equation 26 is very simple and is obtainable directly from volt-time area considerations. However, in many 2-core amplifiers, the evaluation of the signal current needed for a wanted I_L depends on the knowledge of gate-winding currents during saturation intervals. These saturation intervals coincide with the forcing intervals, during which $i_L'=i_g$. This implies that there is need here for calculations of $i_{L\alpha}'$, as given by equation 23. For the sake of convenience, a family of curves expressing c in terms of θ and α is presented in Fig. 6.

In fact, for the evaluation of signal currents it may be of interest in some cases to know the value $i_{L\pi}'$ of load current at the transition from forcing to relaxation. In this situation of continuous flow for the 2-core amplifier, it is found

$$i_{L\pi}'=\frac{V_{gm}}{R_L}\times\left(\frac{\cos\theta\sin\theta-\cos\theta\sin(\alpha-\theta)\epsilon^{-(\pi-\alpha)\cot\theta}}{1-\epsilon^{-\pi\cot\theta}}-k\right)=\frac{V_{gm}}{R_L}(d-k) \qquad (27)$$

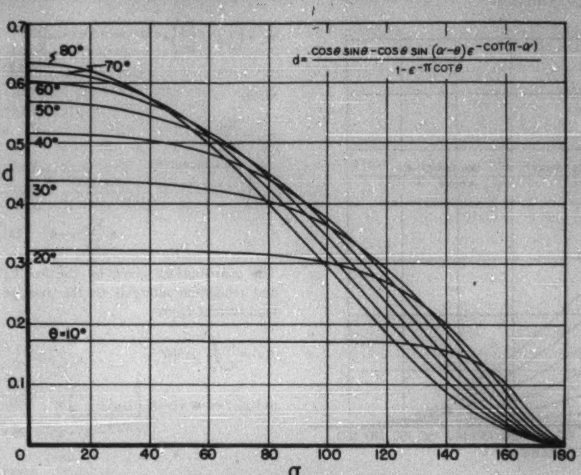


Fig. 7. Quantity d of equation 27

A family of curves expressing d in terms of θ and α is presented in Fig. 7.

5. Two-Core Amplifiers with D-C Output and Load-Discharging Provisions

Situations of this sort occur if the load is supplied by the output of a center-tap amplifier, Fig. 8(C), or also is inserted directly in series with the N_f coils of external feedback amplifiers, Fig. 8(A). In either case, for proper operation, a discharging rectifier is connected directly across the load to offer a relaxation path, while unidirectional current is supplied from the power source to the load during forcing only. All analyses and results of section 4 apply without modifications. The same analysis applies also if the load is in the output circuit of a bridge-type amplifier where a relaxation path short-circuiting the load is offered by the two lower rectifiers in series, Fig. 8(D).

6. Two-Core Amplifiers with D-C Output Without Load-Discharging Provisions

If the load discharging rectifier is removed in the center-tap amplifier of Fig. 8(C), the load current finds relaxation paths through both gate coils N_g. As relaxation initiates, upon reversal of the a-c supply voltage, this voltage is balanced by the rate of change of both core fluxes, while comparatively large currents are induced in the control circuit by transformer action to cancel the magnetomotive forces of the relaxation currents.

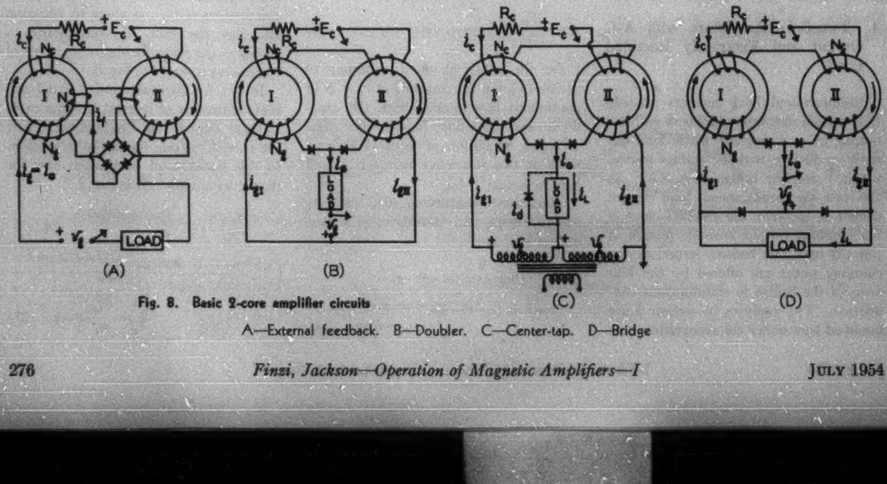
A similar situation occurs in the external feedback of Fig. 8(A) if the load is inserted in the feedback loop without providing a load-discharging rectifier directly across the load. The relaxation current then finds a path within the loop formed by the N_f coils on both cores through the four rectifiers of the bridge in simultaneous conduction, and comparatively large currents again are induced in the control circuit.

The implications of these undesirable types of relaxation in terms of amplifier control are examined in section 8 of the companion paper.[1] But the considerations given in section 4 of the present paper still apply to the determination of the load current for any assumed angle of firing α, at least as long as the resistances of these relaxation paths and reflected circuits do not affect remarkably the relaxation decay.

Conclusions

An analysis of load circuit behavior upon firing is presented for certain types of loads which are of interest in magnetic amplifier applications. This is substantially a treatment of certain linear networks which contain rectifying elements and voltage sources and which are submitted to recurrent switching transients. Nomograms and charts are given which eliminate the need for numerical cut-and-try solutions of transcendental equations appearing in problems of this kind.

For nonrectified loads a quick determination of the angle of extinction β



(A) (B) (C) (D)

Fig. 8. Basic 2-core amplifier circuits

A—External feedback. B—Doubler. C—Center-tap. D—Bridge is made possible in terms of the constants of the circuit and of the angle of firing α taken as the independent variable. Kdowledge of the angle of extinction is indispensable for the evaluation of the output current and for the determination of the saturation interval. (This, in turn, is essential for the further determination of the corresponding amplifier signal requirements.)

For rectified loads means are given to evaluate instantaneous or average load currents separately during forcing and relaxation intervals. (This is essential also for the further determination of the corresponding signal.)

The further question of the evaluation of the signal required for any wanted angle of firing is treated in the companion paper for the same types of loads, with particular reference to certain kinds of 2-core amplifiers with sensitive core materials.

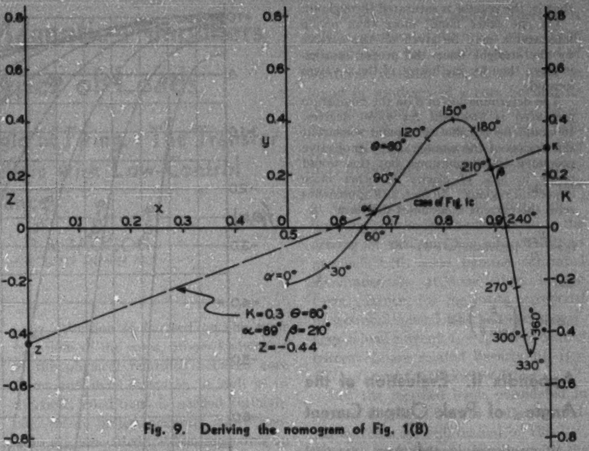


Fig. 9. Deriving the nomogram of Fig. 1(B)

Appendix I. Derivation of Nomogram of Fig. 1(B)

Equation 3 is rewritten

$$[\cos\theta\sin(\alpha-\theta)-k]\epsilon^{\alpha\cot\theta} = [\cos\theta\sin(\beta-\theta)-k]\epsilon^{\beta\cot\theta} \quad (28)$$

where $0 \leq \theta \leq \pi/2,\ 0 \leq \alpha \leq \pi,\ -1 \leq k \leq 1,\ \alpha < \beta$

Or

$$F(\alpha,\theta,k) = F(\beta,\theta,k) = Z \quad (29)$$

and

$$F(\alpha,\theta,k) - Z = 0 \quad (30)$$

$$F(\beta,\theta,k) - Z = 0 \quad (31)$$

Equation 30 can be rewritten as the form

$$\begin{vmatrix} 1 & 0 & Z \\ 0 & 1 & k \\ 1 & \epsilon^{\alpha\cot\theta} & \epsilon^{\alpha\cot\theta}\cos\theta\sin(\alpha-\theta) \end{vmatrix} = 0$$

or, after manipulation

$$\begin{vmatrix} 0 & Z & 1 \\ 1 & k & 1 \\ \dfrac{\epsilon^{\alpha\cot\theta}}{1+\epsilon^{\alpha\cot\theta}} & \dfrac{\epsilon^{\alpha\cot\theta}\cos\theta\sin(\alpha-\theta)}{1+\epsilon^{\alpha\cot\theta}} & 1 \end{vmatrix} = 0 \quad (32)$$

In a system of orthogonal x, y co-ordinates a family of curves for constant θ's is plotted with

$$x = \frac{\epsilon^{\alpha\cot\theta}}{1+\epsilon^{\alpha\cot\theta}}$$

and

$$y = \frac{\epsilon^{\alpha\cot\theta}[\cos\theta\sin(\alpha-\theta)]}{1+\epsilon^{\alpha\cot\theta}}$$

One such curve is shown in Fig. 9. Moreover, a line $x=1$, $y=k$ and a line $x=0$, $y=Z$ are plotted. Thus a straight line through the point α on the θ-curve and through k on the k-axis locates on the Z-axis the value of Z, satisfying equation 30. In view of the identity of equations 30 and 31 the same straight line intersects the θ-curve at point β which satisfies equation 31 for the same θ, k and Z. That is, given α, θ, and k, the corresponding angle of extinction β, defined by equation 28, is found by alignment, drawing a straight line through point α on the proper θ-curve, and the point k on the k-axis. β is determined by the intersection of this line with the same θ-curve, so that the Z-axis is not needed for this determination. Thus no Z-axis is shown in Fig. 1(B), for the purpose of better utilization of the over-all dimensions available for this graph.

It may not be noted that no β intersection may be found for certain sets of α, θ, and negative k values. This is the case when no extinction occurs in the cycle;

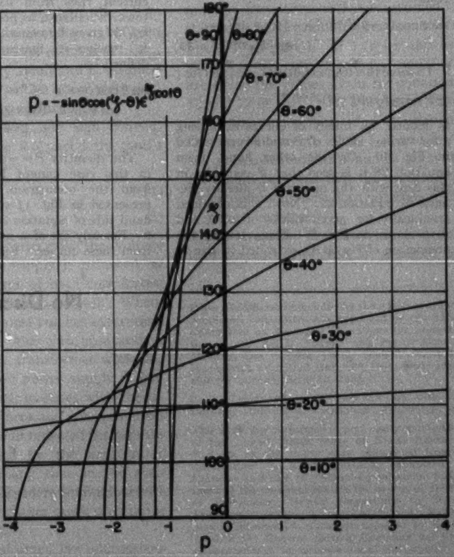


Fig. 10. Quantity P of equation 33

that is, the reactor is saturated throughout.

On the other hand, more than one β intersection may be given on the θ-curve by the straight line; the proper solution is given then by the lowest of the β values located.

The determination of β on the nomogram presented is awkward for $\theta<30$ degrees. It would be possible to expand somewhat this region of the nomogram by projective geometry transformations, but this would result in lack of clarity in other more important regions. In fact, as θ approaches zero, the current i_L assumes the form

$$i_L=\frac{V_{gm}}{R}[\cos\theta\sin(\omega t-\theta)-k]$$

and

$$\beta\cong\sin^{-1}\left(\frac{k}{\cos\theta}\right)+\theta$$

Appendix II. Evaluation of the Angle ζ of Peak Output Current

As mentioned in the paper, practical need is felt in some instances for a determination of the angle ζ at which the amplifier output current reaches its peak. This determination is accomplished by the following means:

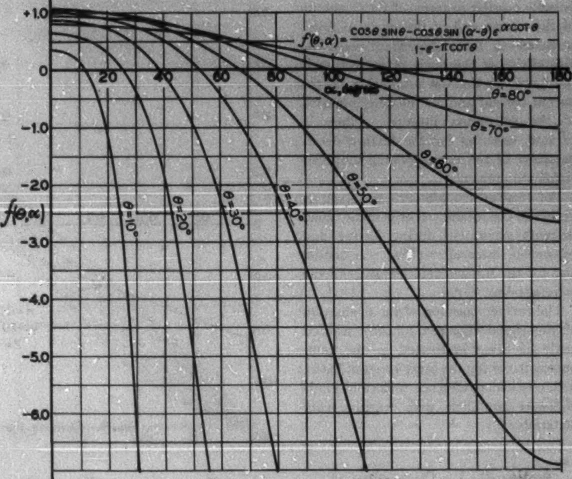


Fig. 11. Right side of equation 34

Case I—Two-Core Amplifiers With A-C Output or With Rectified Output in Discontinuous Flow (Sections 3 through 6)

In either situation the instantaneous load current initiating at α rises toward its crest as per equation 2. By maximizing i_L in this equation, the following relation results

$$\sin\theta\cos(\zeta-\theta)\epsilon^{\zeta\cot\theta}=-[\cos\theta\sin(\alpha-\theta)-k]\epsilon^{\alpha\cot\theta} \quad (33)$$

To solve this transcendental equation for ζ

$$-P=\sin\theta\cos(\zeta-\theta)\epsilon^{\zeta\cot\theta}$$

is defined. A family of curves expressing ζ for various values of P and θ is presented in Fig. 10. On the other hand, from equation 28 it is seen that P coincides in this case with the quantity Z already defined in equation 29. Now, Z is obtained graphically for given values of α, θ, and k, as shown in Fig. 9. Thus, use of the nomogram of Fig. 1(B), extended in size to include the Z-axis in conjunction with the curves of Fig. 10, enables determination of ζ for any set of α, θ, and k.

Case II—Two-Core Amplifiers With Rectified Output in Continuous Flow (Sections 4, 5, and 6)

Upon firing, the instantaneous load current rises from its initial value $i_{L\alpha}'$ towards its crest, as per equation 15, where $i_{L\alpha}'$ is given by equation 23. The current i_L' reaches its maximum at the angle ζ defined by

$$\sin\theta\cos(\zeta-\theta)\epsilon^{\zeta\cot\theta}=\frac{\cos\theta\sin\theta-\cos\theta\sin(\alpha-\theta)\epsilon^{\alpha\cot\theta}}{1-\epsilon^{-\pi\cot\theta}} \quad (34)$$

The quantity $P=-\sin\theta\cos(\zeta-\theta)\epsilon^{\zeta\cot\theta}$ in this case cannot be obtained simply from the nomogram. Thus curves are presented in Fig. 11 expressing the right-hand side of equation 34 in terms of α and θ. For any set of α and θ, P is evaluated from these curves. Use of this value of P in conjunction with the curves of Fig. 10 yields ζ.

References

1. THE OPERATION OF MAGNETIC AMPLIFIERS WITH VARIOUS TYPES OF LOADS. PART II—CONTROLLING THE ANGLE OF FIRING. THE TRANSFER CHARACTERISTICS OF AMPLIFIERS WITH LOW CONTROL IMPEDANCE, **L. A. Finzi, R. R. Jackson.** *AIEE Transactions*, vol. 73, pt. I, July 1954, pp. 279–88.

2. THEORY OF RECTIFIER–D-C MOTOR DRIVE, **E. H. Vedder, K. P. Puchlowski.** *AIEE Transactions*, vol. 62, 1943, pp. 863–70.

3. THE TRANSDUCTOR AMPLIFIER (book), **Ulrik Krabbe.** Lindhska Boktryckeriet. Orebro, Sweden, 1947.

4. ANALYSIS AND DESIGN OF SELF-SATURABLE MAGNETIC AMPLIFIERS, **Sidney B. Cohen.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 39, no. 9, Sept. 1951, pp. 1009–20.

5. TRAITÉ DE NOMOGRAPHIE (book), **M. d'Ocagne.** Gauthier-Villars, Paris, France, 1st ed., 1899.

6. SERIES-CONNECTED MAGNETIC AMPLIFIER WITH INDUCTIVE LOADING, **Thomas G. Wilson.** *AIEE Transactions*, vol. 71, pt. I, Jan. 1952, pp. 101–10.

No Discussion

The Operation of Magnetic Amplifiers with Various Types of Loads

Part II—Controlling the Angle of Firing. The Transfer Characteristics of Amplifiers with Low-Control Impedance

L. A. FINZI
MEMBER AIEE

R. R. JACKSON
ASSOCIATE MEMBER AIEE

THE companion paper (part I)[1] has supplied expressions for output currents and for currents in the load in terms of the angle of firing α taken as the independent variable. The treatment has been developed with reference to practical situations of magnetic amplifier circuitry, but substantially no need was felt for a close understanding of the phenomena taking place in the amplifier proper.

To obtain transfer characteristics relating load currents to amplifier signal a further step must be taken; namely, the signal required to set a wanted angle of firing must be determined. The question is under discussion in the art as to whether the output of a magnetic amplifier is controlled truly by magnetomotive forces (mmf) or by voltages or, conceivably, by some other variable. It may be said that this is a matter of personal taste, as any complete analysis of these electrically and magnetically interlinked systems requires consideration of both Faraday's law and Ampere's circuital law. But certainly for certain types of amplifiers[2] the analysis is attacked profitably in terms of voltage equations only; no need is felt (up to a certain point) to consider the mmf equations of the cores; hence control voltages (or volt-time areas) are taken as the primary factor governing the angle of firing and thus the output. As the analysis in part I supplies information concerning the angle at which a core ceases to be saturated, the evaluation of the signal voltage required for any wanted reset is almost immediate in amplifier circuits of this kind.

On the other hand, the operation of 2-core amplifiers of the kind shown in Fig. 8, part I, is defined best in terms of the mmf of average control current, which is taken as the controlling signal. In the following the question of signal requirements for wanted angles of firing is investigated for conventional amplifiers of this second kind, in situations of low control circuit impedance.

As attention is focused on the mmf equations of the cores, proper knowledge of the physical relations between core fluxes and mmf in static, as well as in dynamic, conditions is needed throughout. More or less accurate approximations are demanded for this, depending on the order of magnitude of the ampere-turn gain of the device.

1. The Transfer Characteristic of the Simple Saturable Reactor

The saturable reactor is an elementary device obtained from the amplifier circuit of Fig. 8 of part I by elimination of the feedback turns N_f and associated bridge. In this case it is acceptable practically to approximate the core magnetization curve by a vertical segment through the origin and two horizontal straight lines. As long as a core flux ϕ is swinging between positive and negative saturation $\pm\phi_{sat}$ its mmf lH is very small and is assumed to be zero; mmf's other than zero yield positive or negative saturation.

With the symbols and sign conventions shown in Fig. 8 of part I the mmf's of the two cores are

$$lH_{I} = N_c i_c + N_g i_g \tag{1}$$

$$lH_{II} = N_c i_c - N_g i_g \tag{2}$$

During prefiring intervals, the voltage of the a-c power supply is balanced by the rates of change of both core fluxes; $lH_I = lH_{II} = 0$, and thus $i_c = i_g = 0$. Whenever one of the cores reaches saturation, firing occurs; the other core, unsaturated, acts as a current transformer while the voltage of the a-c power supply is balanced by the load, in series with the resistances, inductances, and voltage sources of the control circuit reflected through the coupling provided by the unsaturated core. Hence at any time in the cycle $N_c i_c = N_g |i_g|$. (Readers not familiar with this matter may gather quick orientation from reference 3.)

This "d-c transformer" law relating control currents to gate-winding currents is valid, regardless of the tye of load, of the impedance of the control circuit and of whatever residual inductance may be found in windings of a core beyond its sharply defined saturation flux level. Transfer characteristics relating half-cyclic averages of amplifier output mmf $N_g I_g$ to averages of control mmf $N_c I_c$ are represented by portions of straight lines at 45 degrees through the origin, up to breaking points, which correspond to the driving of the amplifier into conditions in which both cores become saturated simultaneously at some time in the cycle. Hence, for the loads considered in section 3, part I, the half-cyclic averages of load current, $I_L = I_g$ and of control current I_c are related linearly by $I_L = I_c N_c / N_g$ over the region $\alpha < \beta < \alpha + \pi$.

For the rectified loads considered in section 4, part I, control and gate-winding currents are still related by this law, but the load current coincides with the gate-winding current only during forcing intervals. As relaxation initiates, i_G drops quickly to zero, while the load current i_L keeps flowing through relaxation paths. That is $I_c N_c / N_g = I_G = I_{LS}'$, but the total half-cyclic load current $I_{LS} = I_{LS}' + I_{LS}''$ where I_{LS}'' is the contribution from the relaxation interval. To calculate load currents for given α it must be known whether the flow of current is discontinuous or continuous, as discussed in sections 2 and 4 of Part I; namely, discontinuous flow is assumed to start with, and δ evaluated from equations 9 and 10 of part I. If $\delta < \alpha + \pi$ this assumption was correct and I_{LS}' and I_{LS}'' are twice the values given by equations 12 and 13 of part I. If δ is greater than $\alpha + \pi$ the assumption of discontinuous flow was incorrect; then I_{LS}' and I_{LS}'' are given by equations 24 and 25 of part I. (It should be noted that the eventuality of discontinuous flow does not arise if E_L is zero or negative.)

The transfer characteristic, relating load currents to signal, is obtained by

---

Paper 54-238, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 23, 1954; made available for printing April 15, 1954.

L. A. FINZI and R. R. JACKSON are with the Carnegie Institute of Technology, Pittsburgh, Pa.

This work was performed in part under sponsorship of the United States Office of Naval Research. Portions of the paper were abstracted from a dissertation by R. R. Jackson, submitted to Carnegie Institute of Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy. Helpful suggestions by F. J. Friedlaender are gratefully acknowledged. Experimental information graciously supplied by H. W. Lord of the General Electric Company has been of assistance in the preparation of the paper.

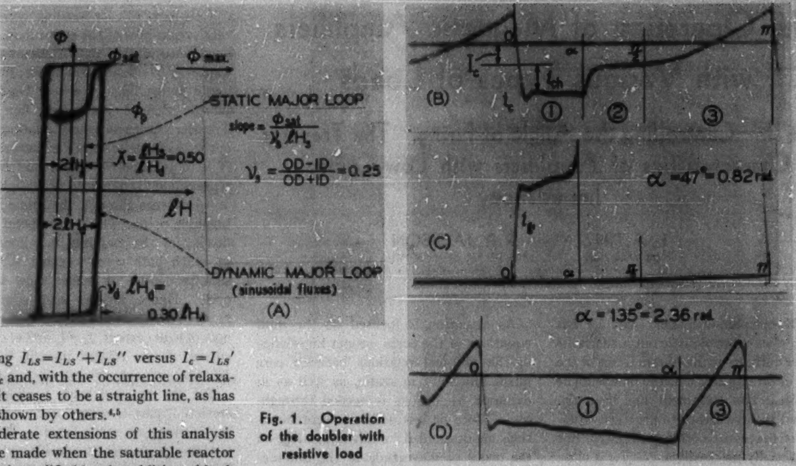


Fig. 1. Operation of the doubler with resistive load

plotting $I_{LS}=I_{LS}'+I_{LS}''$ versus $I_c=I_{LS}' N_g/N_c$ and, with the occurrence of relaxation, it ceases to be a straight line, as has been shown by others.[4,5]

Moderate extensions of this analysis can be made when the saturable reactor circuit is modified by the addition of feedback turns N_f shown in Fig. 8(A), part I. The mmf equations of the two cores become

$$lH_I = N_c i_c + N_g i_g + N_f i_f \qquad (3)$$

$$lH_{II} = N_c i_c - N_g i_g + N_f i_f \qquad (4)$$

Under the assumption $i_f = |i_g|$ (i.e., perfect rectifiers and no bridge overlap phenomena) these equations show that the d-c transformer law modifies into $N_c i_c = N_g |i_g| (1 \mp N_f/N_g)$, where the sign $-$ or $+$ applies, depending on whether the mmf of feedback and of signal are concordant or discordant. Transfer characteristics relating half-cyclic averages of signal and gate-winding currents remain straight lines (properly tilted) regardless of the type of load. If relaxation occurs in the load, half-cyclic averages of load currents I_{LS} are calculated as before. Transfer characteristics are obtained by plotting $I_{LS}=I_{LS}'+I_{LS}''$ versus $I_c = I_{LS}'(1 \mp N_f/N_g)N_g/N_c$.

However, this oversimplified analysis loses practical validity as the number of feedback turns N_f is increased, and becomes absurd when $N_f = N_g$. This inadequacy is disturbing because conventional self-saturating amplifiers of the doubler, center-tap or bridge types (most common in the manufacturing practice) behave exactly as external feedback amplifiers with $N_f = N_g$. A proof of this equivalence results from simple algebraic transformations of the corresponding circuit equations.[6]

Attempts to improve on this analysis by introducing a finite slope for the unsaturated portion of the magnetic characteristic are only moderately useful and are inherently artificial. A more realistic and valuable approach has been chosen by Storm[7] by discarding single-valued representations and by expressing the magnetization curve in the unsaturated region by a rectangle with vertical sides. This dynamic hysteresis loop yields the mmf of the amplifier cores (and thus the instantaneous control currents) during the prefiring interval of swinging fluxes.

But no information is expressed in this way on the mmf requirement of the unsaturated core during the remaining part of the half-cycle. If the static hysteresis loop is very narrow its mmf during this part can be assumed to be zero. But this simplification is less adequate for common cores (i.e., *0.002*-inch Deltamax, Hypernik *V*, etc.).[8,9] In fact, during the output interval, while the one core is definitely saturated, the other, unsaturated, wanders slowly along the bottom of the minor loop from the left toward the right side, and the instantaneous control current is not zero nor does it average zero over this interval. The time-varying mmf of the unsaturated core is governed by a variety of complex phenomena presently understood only in part. Further analytical attempts must be preceded by a careful and discriminating observation of such elusive effects.

Semiquantitative considerations follow, based on experimental investigations of this kind, together with circuit analysis, to remove some of the vagueness which affects the considerable empirical corrections (or "shearing" factors) commonly needed to match analytical predictions to actual experimental transfer characteristics. Particular attention is given to the positive (high-gain) branch of the transfer characteristic. Purely resistive loads are considered first, then extensions are made to more general cases.

2. The Transfer Characteristics of the Doubler with Resistive Loads

The amplifier circuit of Fig. 8(B), part I, is considered, with load R_L, purely resistive. The mmf equations, with the symbols and sign conventions shown, are

$$lH_I = N_c i_c + N_g i_{gI} \qquad (5)$$

$$lH_{II} = N_c i_c + N_g i_{gII} \qquad (6)$$

and the voltage equations are, for the gate loops

$$v_g = V_{gm} \sin \omega t = N_g \frac{d\phi_I}{dt} + r_g i_{gI} + R_L i_G + v_{\gamma I} \qquad (7)$$

where $i_G = i_{gI} - i_{gII}$

$$v_g = V_{gm} \sin \omega t = -N_g \frac{d\phi_{II}}{dt} - r_g i_{gII} + R_L i_G - v_{\gamma II} \qquad (8)$$

where r_g is the ohmic resistance of one gate winding and $v_{\gamma I}$ and $v_{\gamma II}$ are the voltages across the rectifiers (positive drops for the chosen positive direction of currents). Moreover

$$E_c = R_c i_c + N_c \frac{d\phi_I}{dt} + N_c \frac{d\phi_{II}}{dt} \qquad (9)$$

for the control loop where E_c is the d-c voltage applied, R_c is the total resistance of the control circuit.

Defining

$$i_c=I_c+i_{ch} \tag{10}$$

(where $I_c = E_c/R_c$ and i_{ch} is the harmonic content of the control current, averaging zero over the half-cycle) equation 9 transforms into

$$0=R_c i_{ch}+N_c\frac{d\phi_I}{dt}+N_c\frac{d\phi_{II}}{dt} \tag{11}$$

This algebraic transformation corresponds to a substitution of the actual control circuit with two separate circuits, both with N_c turns on each core. One circuit is supplied from a constant current source I_c (where $I_c=E_c/R_c$ by definition); the other, with total resistance R_c, has no applied voltage, i.e., is short-circuited.

By subtracting equation 8 from 7

$$0=N_g\frac{d\phi_I}{dt}+N_g\frac{d\phi_{II}}{dt}+r_g i_{gI}+r_g i_{gII}+v_{\gamma I}+v_{\gamma II} \tag{12}$$

which can also be written directly as the voltage equation of the external loop of the two gate windings in series. Using equation 11, equation 12 transforms into

$$0=r_g i_{gI}+r_g i_{gII}+v_{\gamma I}+v_{\gamma II}-R_c i_{ch}N_g/N_c \tag{13}$$

This equation is useful as it gives some clue as to which rectifier is conducting at any time.

It is assumed at first that the rectifiers have no appreciable leakage; that is, leakage, if any, shall be considered later by way of corrections; see section 7. Rectifiers satisfying this condition are readily available to the experimenter. Also it is assumed that the two cores have equal sizes and identical magnetic properties, as should be the case if proper core-matching techniques are used in amplifier production.

The customary assumption $R_c=0$ is avoided here because of possible misunderstandings about its meaning. Rather, in the joint experimental and analytical study that follows, cases of low control circuit resistance and cases in which this resistance is increased, e.g., by the addition of sizable forcing resistors, are considered in succession.

The operation of a doubler with purely resistive loads is evidenced by the oscillograms of Fig. 1. (Core data: Hypernik V, inside diameter (ID) 1$^1/_2$ inches; outside diameter (OD) 2$^1/_2$ inches; height 1/2-inch, thickness of laminations 0.002 inch.) In this case R_c (or more properly R_c/N_c^2) is low, as the control winding on each core occupies about one-half the available winding space, while the control voltage source E_c has negligible internal resistance. For this amplifier the rectifiers are Germanium *AJA1A3*, General Electric Company, $N_c=N_g=2{,}500$ turns, $R_c=2\times 65$ ohms, $r_g=75$ ohms. The load resistance is $R_L=400$ ohms in the tests of Fig. 1; moreover, the crest V_{gm} of the 60-cycle gate supply voltage is $V_{gm}=\sqrt{2}\times 130$ volts; this value is only slightly lower than the crest $V_{gN}=\omega N_g\Phi_{sat}$ of the "normal-excitation" voltage, swinging the core fluxes from positive to negative flux saturation level $\pm\Phi_{sat}$, and vice versa.

Fig. 1(A) shows the major dynamic loop of one core at the operating 60-cycle frequency and a minor loop described in the operation at a rather early angle of firing over the positive range of the transfer characteristic; Fig. 1(B) and Fig. 1(C) show instantaneous values of i_c and i_{gI} (at the same scale) during one positive half-cycle. (During this half-cycle no detectable i_{gII} is recorded.)

INTERVAL 1

At $\omega t=0$, core I is found at some initial (preset) flux level Φ_p, while core II is at a level somewhere on the top of the loop near the residual level Φ_{res}. During the prefiring interval Φ_I climbs toward $+\Phi_{sat}$ while Φ_{II} diminishes at a nearly equal and opposite rate, as shown by equation 11. Both fluxes are varying cosinusoidally in time, or very nearly so, and their mmf requirements result from portions of a minor loop, which in its descending branch is seen to be very nearly coincident with the major dynamic loop embodying effects of eddy currents and domain boundary movements.

During this interval i_{gII} cannot be negative in the absence of rectifier leakage. It is shown here that it cannot be positive either. In fact, $r_g i_{gI}$ and $v_{\gamma I}$ are certainly positive, hence equation 13 indicates that the sum $r_g i_{gII}+v_{\gamma II}-R_c i_{ch}N_g/N_c$ must be negative. Now over the positive branch of the transfer characteristic $|i_c|>|I_c|$ in the prefiring interval from 0 to α, as I_c is obtained from an averaging process of i_c over the whole half-cycle. That is, with our sign conventions, i_{ch} is a negative quantity during prefiring and thus $-R_c i_{ch}N_g/N_c$ is a positive number. Since i_{gII} cannot be negative, it follows that $v_{\gamma II}$ must be negative; i.e., a reverse voltage appears across the gate rectifier II and therefore i_{gII} cannot be positive either.

(On the basis of superficial thinking it might be taken for granted that no positive i_{gII} can flow during the half-cycle considered. Such intuitive reasoning will be shown to be fallacious. Positive i_{gII} may well flow in later portions of the half-cycle considered).

With $i_{gII}=0$, equation 6 yields $i_c=lH_{II}/N_c$ where lH_{II} is obtained from the descending branch of the major dynamic loop. Analytically this loop can be approximated by a parallelogram of width $2lH_d$ with slopes $d\Phi/dlH=\Phi_{sat}/\nu_d lH_d=$ constant. Since equation 8 yields

$$\Phi_{II}=\Phi_{sat}-(1-\cos\omega t)V_{gm}/N_g\omega$$

$$lH_{II}=-lH_d\{1-\nu_d+\nu_d V_{gm}/V_{gN}-\nu_d V_{gm}/V_{gN}\cos\omega t\} \tag{14}$$

where $V_{gN}=\omega N_g\Phi_{sat}$

Hence the contribution I_c given by this first prefiring interval to the total I_c is

$$I_c'=\frac{1}{\pi}\int_0^\alpha i_c d\omega t=\frac{1}{\pi N_c}\int_0^\alpha lH_{II}d\omega t$$
$$=-\frac{lH_d}{\pi N_c}\left\{\left(1-\nu_d+\nu_d\frac{V_{gm}}{V_{gN}}\right)\times\alpha-\nu_d\frac{V_{gm}}{V_{gN}}\sin\alpha\right\} \tag{15}$$

which for $V_{gm}\cong V_{gN}$ reduces to

$$I_c'=-\frac{lH_d}{\pi N_c}(\alpha-\nu_d\sin\alpha) \tag{16}$$

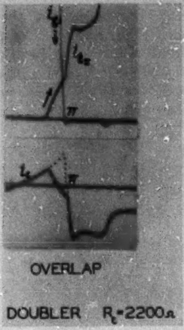
The simplified equation 16, coincides with an expression derived by Phillips[9] but it is only a part of the total I_c. (The assumption $V_{gm}\cong V_{gN}$ is accepted throughout the following text.)

INTERVAL 2

At $\omega t=\alpha$ the ascending Φ_I reaches $+\Phi_{sat}$ and from there on $d\Phi_I/dt$ takes an entirely different (and much smaller) order of magnitude; according to equation 7, i_{gI} is now limited primarily by the load and gate-winding resistances. (The phenomena taking place during the firing transition are not easily described; the firing shock may cause some oscillations in the various circuits with spikes and dips in i_c. Fortunately these phenomena are short-lived and may be ignored in the evaluation of I_c by an averaging integration extended over the whole half-cycle.)

If firing occurs at $\alpha<\pi/2$, i_{gI}, after its firing jump, keeps increasing, reaching its peak $i_{gI\,peak}=V_{gm}/R$ at $\omega t\cong\pi/2$, where $R=R_L+r_g$. Accordingly Φ_I at that time reaches its maximum value Φ_{max}, where $\Phi_{max}>+\Phi_{sat}$ since the saturated incremental permeabilities of core I, small as they appear, are certainly not zero.

Equation 11 indicates that Φ_{II} keeps decreasing during this second time interval. (This is generally true because i_{ch} is positive over this interval for all but very small values of α; also $R_c i_{ch}$ is generally small in comparison to $N_c d\Phi_I/dt$ over this interval). But both flux changes are now slower, by an entirely different order of magnitude. Therefore the mmf requirements of the still descending core II



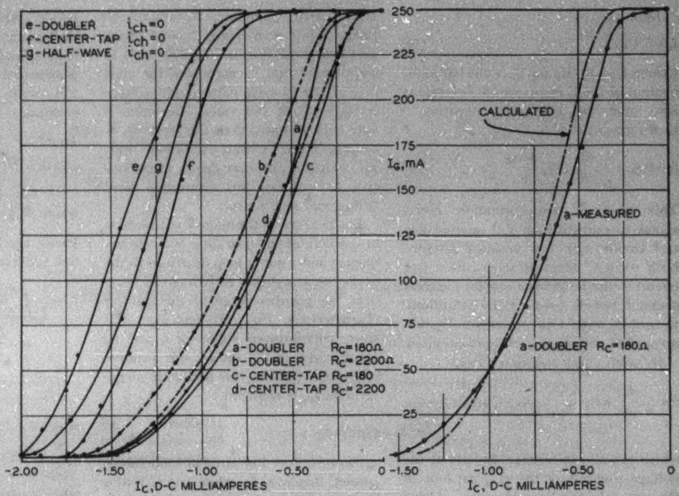


Fig. 2. Transfer characteristics, resistive load. Influence of forcing control resistors

are no longer expressed by its dynamic major loop and are given more nearly by its static major loop. This loop can be approximated by a parallelogram of width $2lH_s$ and of slope $d\Phi/dlH=\Phi_{sat}/\nu_s lH_s$ where $\nu_s=(\mathrm{OD}-\mathrm{ID})/(\mathrm{OD}+\mathrm{ID})$.[11]

Considering that the downward excursion on the static loop is very limited, it appears that i_c over this interval (in which i_{gII} is still zero) is expressed by

$$i_c=-\chi\frac{lH_d}{N_c}(1-\nu_s\cos\alpha)=\text{constant}$$

where $\chi=2lH_s/2lH_d$

The contribution I_c'' given to the total I_c by this interval between α and π is

$$I_c''=-\chi\frac{lH_d}{\pi N_c}(1-\nu_s\cos\alpha)(\pi/2-\alpha) \qquad (17)$$

valid for $\alpha<\pi/2$

No such contribution appears in the overall expression of I_c, if firing occurs at $\alpha\geqslant\pi/2$ (as evidenced by Fig. 1(D), which shows the instantaneous control current for a firing angle of 135 degrees). Thus

$$I_c''=0 \text{ if } \alpha\geqslant\pi/2 \qquad (18)$$

INTERVAL 3

In the remaining part of the half-cycle i_{g1} decreases and accordingly Φ_I also decreases from Φ_{max}. Hence a process begins in which, with $R_c i_{ch}$ small in comparison to $|N_c d\Phi_I/dt|$, the flux Φ_{II} starts increasing, very slowly, toward the level $\Phi_{II}=\Phi_p$ at the end of the half-cycle. During this third time interval the mmf requirements of the unsaturated core II wandering from the left to the right side of the static loop are somewhat obscure.

Attempts to give analytical expressions to the incremental permeabilities corresponding to this portion of the bottom of the minor loop have not been satisfactory so far, as any approximation tried does not explain some intriguing experimental evidence; for instance, the oscillographic trace of lH_{II} as a function of time over this third interval seems to be affected only slightly by changes of R_L or V_{gm} (and thus of $i_{g\ peak}$) or of R_c within somewhat broad ranges. (This is a matter for further investigation and analyses presently in course).

But the very invariance of this trace, recognized from experimentation performed on various amplifiers for different cores, allows for an empirical approximation of i_c over this third interval. Namely, with $i_{gII}=0$ (as is still well the case for low values of R_c) the parabolic approximations are suggested

$$i_c=-\frac{\chi lH_d}{N_c}\left\{\frac{2\omega t(\pi-\omega t)}{(\pi-\pi/2)\pi/2}-(1-\nu_s\cos\alpha)\right\}$$
$$=-\frac{\chi lH_d}{N_c}\left\{8\left[\frac{\omega t}{\pi}-\left(\frac{\omega t}{\pi}\right)^2\right]-(1-\nu_s\cos\alpha)\right\} \text{ if } \alpha\leq\pi/2 \qquad (19)$$

or

$$i_c=-\frac{\chi lH_d}{N_c}\left\{\frac{2\omega t(\pi-\omega t)}{\alpha(\pi-\alpha)}-(1-\nu_s\cos\alpha)\right\} \text{ if } \alpha\geq\pi/2 \qquad (20)$$

Hence the contribution given by i_c over this third interval is

$$I_c'''=\frac{1}{\pi}\int_{\pi/2}^{\pi}i_c d\omega t=-\frac{\chi lH_d}{2N_c}\times(1/2-\nu_s\cos\alpha) \text{ if } \alpha\leq\pi/2 \qquad (21)$$

or

$$I_c'''=\frac{1}{\pi}\int_{\alpha}^{\pi}i_c d\omega t=-\frac{\chi lH_d}{N_c}\left(\frac{\pi-\alpha}{\pi}\right)\times(\pi/3\alpha-1/3-\nu_s\cos\alpha) \text{ if } \alpha\geq\pi/2 \qquad (22)$$

The total control current I_c needed for any wanted angle of firing is thus

$$I_c=I_c'+I_c''+I_c'''=\text{function of } \alpha \qquad (23)$$

where I_c' is given by equation 16 while I_c'' is given by equation 17 or equation 18 and I_c''' by equation 21 or equation 22 depending on whether $\alpha\leq\pi/2$ or $\alpha\geq\pi/2$.

On the other hand the load current for given α is obtained from the type of reasoning used in part I, with some correction for contributions given by prefiring gate currents, whenever significant, i.e.

$$I_L=I_G=I_{g1s}+I_{g1p}=\text{function of } \alpha \qquad (24)$$

where

$$I_{g1s}=\frac{1}{\pi}\int_{\alpha}^{\pi}i_{g1}d\omega t=\frac{V_{gm}}{\pi R}(1+\cos\alpha), \quad (\text{with } R=R_L+r_g) \qquad (25)$$

is the contribution given to the half-cyclic average of gate current by the most significant saturation (postfiring) interval, and

$$I_{g1p}=\frac{1}{\pi}\int_0^{\alpha}i_{g1}d\omega t\cong 2\frac{lH_d}{N_g}\frac{\alpha}{\pi} \qquad (26)$$

is the contribution of the magnetization (prefiring) part of the half-cycle. Equation 26 is based on the coarse assumption that the mmf requirement of core I climbing toward saturation is obtained from $lH_1 = N_c i_c + N_g i_g$ with $lH_1 = +lH_d$ throughout. This approximation yields most nearly correct results for $\alpha \cong \pi$ where I_{gxp} may be most significant, while its significance becomes per centwise smaller as α approaches zero.

Under elimination of α from the two parametric equations 23 and 24 the transfer characteristic $I_L = f(I_c)$ is obtained. Also an analytical expression is obtainable for the incremental ampere-turn gain K_{AT}.

$$K_{AT} = \frac{dI_L N_g}{dI_c N_c} = \frac{N_g}{N_c} \frac{dI_L/d\alpha}{dI_c/d\alpha} = \text{function of } \alpha \tag{27}$$

Equations 19 and 20 have been formulated carefully to meet the physical requirement that the incremental ampere-turn gain K_{AT} be continuous at $\alpha = \pi/2$ regardless of whether this angle of firing is reached from $\alpha < \pi/2$ or $\alpha > \pi/2$.

Curve *a* in Fig. 2 shows an experimental transfer characteristic for the amplifier and load described. The same curve is repeated on the right and compared with a calculated one

It should be clear at this point that the equations suggested are only attempts to co-ordinate analytically certain phenomena. The considerations given may be of some value insofar as phenomena are pointed out which should be properly considered and weighed by designers searching for equations of their own. The relative significance of the constants H_d, H_s, ν_d, and ν_s is recognized by inspection of the equations. This may be of interest in the question of pertinent core-testing procedures.

The operation described is not limited to cases of finite positive slopes of the major loop. In fact, $d\Phi/dH$ could be infinite or even negative (as mentioned by Phillips[9]) and no immediate reason could be seen in this to decide upon instabilities in the operation. Whenever the dynamic loop cannot be approximated completely by a parallelogram as it presents bulges, e.g., at the beginning or end of the descending branch, appropriate modifications are easily introduced in the expression of I_c'. No attempts have been made to embody in the equations triggering effects, recognized generally at the lower end of the characteristics and discussed in reference 12.

INFLUENCE OF FORCING RESISTORS

The assumption of very low control circuit resistance can be removed now, and examination may be made as to how the control current i_c is modified during the three intervals described when R_c is increased within certain limits by means of forcing resistors added in the control circuit. Equation 11 shows now that the rates of change of Φ_I and Φ_{II} can no longer be considered as equal and opposite, though for practical values of the forcing resistor, their orders of magnitude in the various intervals may not be affected substantially.

During the prefiring interval i_{ch} is mostly negative, then $|d\Phi_{II}/dt| < d\Phi_I/dt$, thus $v_{\gamma II} < 0$ and $i_{gII} = 0$. Φ_{II} descends at a rate which has been slowed down by the presence of the forcing resistor, but as long as $R_c i_{ch} N_g/N_c$ is small in comparison to v_g, this rate of change remains of an order of magnitude that justifies the use of the major dynamic loop to obtain i_c (and thus I_c') during this interval.

Again, at $\omega t = \alpha$, i_{gI} jumps and, if $\alpha < \pi/2$, it keeps increasing further until it reaches its maximum at $\omega t = \pi/2$ or very nearly so. During this interval, as Φ_I keeps increasing slowly, Φ_{II} decreases, also slowly, and the static loop supplies i_c. No general statements can be made about the sign of i_{ch} and therefore about the influence of the voltage $R_c i_{ch}$ upon the rate of change of Φ_{II} during this interval. For angles of firing already a little larger than 0, i_{ch} is generally slightly positive; then the forcing resistor increases somewhat the rate of change of the very slowly descending Φ_{II}, and in fact Φ_{II} may keep descending along the static loop even while i_{gI}, decreasing from its peak, causes Φ_I to recede from $\Phi_{\max}$. In this case the second interval is protracted beyond $\pi/2$ and this results ultimately in a somewhat larger value of $|I_c|$ for given α.

Other more striking effects, which take place in the doubler, generally toward the end of the voltage half-cycle, are more largely responsible for the over-all increase of $|I_c|$ owing to the forcing resistors. In fact, as core II is wandering from the left to the right side of the static loop the unexplained wave form of lH_{II} appears substantially unaffected by the introduction of the forcing resistor. But already before the reversal of line voltage a positive flow of i_{gII} (with the sign conventions of Fig. 8(B), part I) can be observed. In fact, equation 13 with $r_g i_{gI}$ approaching zero and with i_{ch} certainly positive over this region, indicates the possibility that $v_{\gamma II}$ may no longer be negative.

This is an "overlap" phenomenon of simultaneous conduction of both rectifiers, which is well evidenced in Fig. 2. As i_{gII} now gives a (positive) contribution to lH_{II}, i_c is modified. An exact analysis of the overlap does not, for obvious reasons, seem easy; all the various terms appearing in equation 13 are of the same order of magnitude (and very small) as the end of the half-cycle nears; departures of i_{gI} from the sinusoidal wave forms are noticed; threshold voltages of both rectifiers are significant and, moreover, while one recognizes that i_{ch} is positive now, no quantitative expressions for it are available. Qualitatively one sees that overlap initiates earlier if $R_c N_g^2/N_c^2$ is large in comparison to the gate-winding resistance r_g, and if the threshold voltages of the rectifiers are low.

The transfer characteristic *b* in Fig. 2 evidences the increase of control current requirements due to a sizable increase of R_c. On the other hand, curve *e* indicates the effect of total constraint when i_{ch} is made zero by the use of a 4,500-henry reactor (3.4 megohms for the second harmonic of the 60-cycle power supply frequency) in series in the control circuit.

Finally, curve *g* shows the characteristic obtained multiplying by 2 the output of a half-wave amplifier with constant control current, $i_c = I_c$. The preceding description of the phenomena taking place in the unconstrained doubler indicated already that hasty generalizations of constrained half-wave amplifier behavior to unconstrained doublers, popular as they may be, are not warranted, unless one is satisfied with the rather obvious recognition that in either case signal currents somehow control the amplifier between the two limits of full and minimum output.

It should be added that additional mmf introduced by means of bias windings simply translate horizontally the characteristic *a*. On the other hand, the characteristic obtained in *b* is transformed into *a*, and then translated horizontally, if the described effects of the comparatively high-constraining R_c used to obtain *b* are eliminated by the parallel action of some low resistance bias circuit. More precisely, in the case of many control or bias circuits excited by d-c voltages, the output is dictated by the algebraic sum of their average mmf forces. The degree of constraint of the amplifier, and thus the effects of overlap, etc., depend on the magnitude of the summation $\Sigma N^2/R$ extended to all control circuits.

THE NEGATIVE BRANCH OF THE CHARACTERISTIC

These considerations apply over the so-called positive branch (i.e., the high-gain portion) of the transfer characteristic, in which the output is controlled from its maximum down to its minimum value by increases of the absolute value of the negative I_c. Minimum ouput is obtained when $\alpha=\pi$, that is, when $I_c=-lH_d/N_c$. In this particular condition no firing occurs, as both fluxes swing up and down without ever exceeding the levels $\pm\Phi_{sat}$, and the cyclic average of output current $I_L=I_G$ is given by $I_G=I_{g1p}\cong 2|I_c|N_c/N_g$.

Further increases of $|I_c|$ cause the output to rise again, though with considerably smaller ampere-turn gains. To describe this branch of the characteristic, let it be assumed that at the beginning of the positive voltage half-cycle Φ_I is at some level near the negative residual flux and that Φ_{II} is at some level between the lower and upper limits of the major dynamic loop. At $\omega t\cong 0$ the usual process of swinging fluxes initiates as Φ_I climbs and Φ_{II} comes down. The mmf requirements of the descending core may still be dictated by the major dynamic loop, but this may not be sufficient to define i_c in this case, because with $|I_c|>lH_d/N_c$, the harmonic term i_{ch} is positive and overlap (that is, flow of positive i_{gII}) is observed frequently in the pre-firing interval. Higher values are taken by i_c after firing. In fact, firing occurs here when, at some $\omega t=\alpha$, the descending flux Φ_{II} reaches $-\Phi_{sat}$. From then on the a-c source sees the load resistance, in series with the reflected control circuit, since core I is not saturated and acts as a current transformer. During the firing interval $i_c\cong -i_{gI}N_g/N_c$. In conclusion, the overall ampere-turn gain N_gI_G/N_cI_c, (which at $\alpha=\pi$ was seen to be equal to two) decays gradually, to become very nearly unity when $|I_c|$ is increased to give maximum output ($\alpha=0$) at the left hand of this branch. In view of the comparatively limited significance of this branch of the transfer characteristic a more detailed analysis is omitted.

3. The Center-Tap and Bridge Circuits with Resistive Load

Without writing the voltage and mmf equations of these circuits, it can be stated that the phenomena described in section 2 are recognized without appreciable differences in these new cases. In fact, oscillograms of gate and control current are quite similar to those presented in Fig. 1, but for the presence of some small rectifier leakage. (Of course, the higher reverse rectifier voltages appearing during saturation intervals of these circuits require the use of differently rated rectifying elements.) Equations 15 to 23 apply here too, separate corrections to account for leakage can be introduced subsequently, as mentioned in section 7.

As R_c is increased by means of forcing resistors the first of the two effects previously mentioned in the third interval still appears, that is, i_c remains at the level dictated by the static loop beyond the point at which i_{gI} reaches its peak. However, the other effect (overlap) is much less conspicuous and hardly noticeable. In fact, the voltages to be overcome to initiate any earlier flow of positive i_{gII} are inherently higher in these two amplifiers. As overlap is more nearly negligible, it may be expected that the transfer characteristics of these circuits should be much less sensitive to moderate increases of R_c than was noticed in the doubler. This is evidenced by a comparison of the experimental transfer characteristics *c* and *d* of Fig. 2. An entirely different situation arises, however, in the case of extreme constraint, as shown by curve *f* in Fig. 2.

4. The Doubler with Nonrectified Inductive Loads

This kind of load has been treated in section 3, part I. For any value of α the nomogram supplies the corresponding angle of extinction β. Interval 1 initiates here upon extinction of the preceding saturation interval. That is, at $\omega t=\beta-\pi$ Φ_I is found at the preset level Φ_p and starts climbing toward saturation, while Φ_{II} decreases at a nearly equal rate from its initial level, somewhere near Φ_{res}. The rates of change of these fluxes may be somewhat different from those in the pre-firing interval of the situation described in section 2, because a different portion of the sine wave of v_g is being absorbed by the cores. By way of approximation, however, it shall be assumed that the portion of minor loop described by the descending core over this interval still coincides with a portion of the major dynamic loop. Thus, for this interval, as $\Phi_{II}=\Phi_{sat}-[\cos(\pi-\beta)-\cos\omega t]V_{gm}/\omega N_g$ with the major dynamic loop approximated by a parallelogram, (and with $V_{gm}\cong V_{gN}$) it is found

$$I_c'=\frac{1}{\pi}\int_{\beta-\pi}^{\alpha} i_c d\omega t=-\frac{lH_d}{\pi N_c}\{[1-\nu_d(1+\cos\beta)](\alpha-\beta+\pi)-\nu_d(\sin\alpha+\sin\beta)\} \quad (28)$$

At $\omega t=\alpha$ firing takes place and i_{gI} starts increasing; thus in this second interval ϕ_I increases from ϕ_{sat} toward $\Phi_{max}>\Phi_{sat}$ and accordingly Φ_{II} comes down at a very slow rate so that the mmf requirements of core II are more nearly expressed by the static major loop. The downward excursion on this loop is very limited, hence i_c is substantially constant during this process and is given by

$$i_c=-\frac{\chi lH_d}{N_c}[1-\nu_s(1+\cos\alpha+\cos\beta)] \quad (29)$$

This second interval lasts until i_{gI} reaches its peak at $\omega t=\zeta$, where ζ is evaluated from Appendix II and Fig. 10 of part I. Hence

$$I_c''=\frac{1}{\pi}\int_{\alpha}^{\zeta} i_c d\omega t=-\frac{\chi lH_d}{\pi N_c}\times [1-\nu_s(1+\cos\alpha+\cos\beta)](\zeta-\alpha) \quad (30)$$

Beyond ζ the current i_{gI} decreases from its peak, Φ_I decreases accordingly, and the third interval takes place, during which Φ_{II} rises slowly while the core mmf wanders from the left to the right side of the static major loop. Again the wave form of lH_{II} can be approximated by a parabola. Hence

$$I_c'''=\frac{1}{\pi}\int_{\zeta}^{\beta} i_c d\omega t=-\frac{\chi lH_d}{\pi N_c}\times [{}^1/_3-\nu_s(1+\cos\alpha+\cos\beta)](\beta-\zeta) \quad (31)$$

It is noted that all three intervals are always present for this kind of load. Regardless of the value of α, the post-firing i_{gI} shows a gradual rising in the interval from α to ζ; this is the interval 2. (It is seen here that precise knowledge of ζ is not very essential, as the transition from interval 2 to interval 3 is very gradual in terms of i_c.)

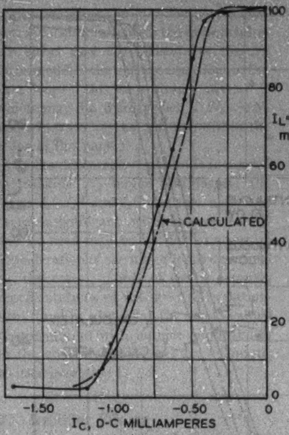
Fig. 3 shows oscillograms of i_c and i_{gI} to different scales. Also transfer characteristics with $R_c=180$ ohms are shown. Here $R_L=584$ ohms, $L=1.86$ henrys, and a d-c source with $K=+0.20$ is included within a rectifying bridge of its own, as shown in Fig. 4, part I.

These considerations apply for low values of R_c. If forcing resistors are used, both effects of extensions of the second interval and of overlap before reaching β are observed again, with resulting increases of the half-cyclic control current requirements for given output.

5. Rectified Loads with Discharging Paths

In these cases, treated in sections 4 and 5 of part I, the load current i_L coincides with the amplifier output current i_G during forcing intervals; the relaxation current i_L'' does not flow through the gate windings and therefore does not affect the control requirements for given α.

If it is known that the flow of load current is discontinuous, ζ is obtained again from part I, Appendix II, Fig. 10. In more usual cases of continuous flow, ζ is found from part I, Appendix II, Fig. 11, in conjunction with Fig. 10.



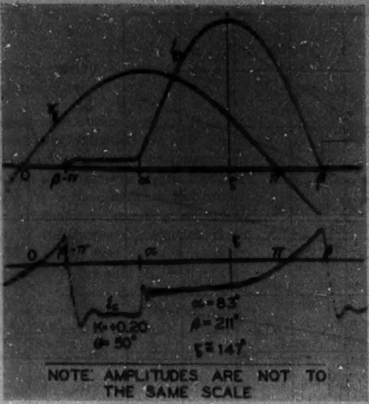


Fig. 3. A-c output. Nonrectified inductive load

After $\omega t = \zeta$, i_{g1} (and Φ_1) decreases; Φ_{II} starts rising and a third interval initiates in which the mmf of the unsaturated core wanders from the left toward the right side of the static loop. In the analysis of part I, sections 2 and 4, no particular effort was made to evaluate the decay of gate current past the angle $\gamma \cong \pi$ at which relaxation initiates and the load current becomes independent of the gate-winding current. On the other hand, it is seen now that this third interval really terminates when the gate current extinguishes; because of some saturated inductance L_g of the gate winding, this occurs at some angle $\eta > \pi$, and the current i_{g1} over this interval is defined by the differential equation 7 of part I. The angle η is found from the curves of Fig. 4, in terms of the initial value i_π of gate current when relaxation initiates (or, actually in terms of $i_\pi R/V_{gm}$), and of $\theta_G = \tan^{-1}(\omega L_g/r_g)$ (i_π is obtained for part I, equation 9 for cases of discontinuous flow, and from part I, equation 27 for the more common continuous flow).

Under these considerations the term I_c' is obtained from integration of

$$i_c = -\frac{lH_d}{N_c}\{1 - \nu_d(1 + \cos\eta + \cos\omega t)\}$$

from $\eta - \pi$ to α

Likewise I_c'' is obtained from integration of

$$i_c = -\frac{xlH_d}{N_c}\{1 - \nu_s(1 + \cos\eta + \cos\alpha)\}$$

from α to ζ

Careful observation of oscillograms of the mmf of the unsaturated core shows that interval 3 consists of two subintervals, from ζ to π and from π to η, with a break in the slope at π. But to obtain workable formulas the whole course of interval 3 can be justifiably approximated by a single parabolic expression

$$i_c = -\frac{xlH_d}{N_c}\left\{1 - \nu_s(1 + \cos\eta + \cos\alpha) - 2\left(\frac{\omega t - \zeta}{\eta - \zeta}\right)^2\right\} \text{ valid from } \zeta \text{ to } \eta$$

(All this is a matter of refinements which cannot be carried out further, because L_g is not very well defined numerically, as the process of decaying i_{g1} extends over a region in which transitions from $+\Phi_{sat}$ to Φ_{res} are taking place in the saturated core.) Tentatively the following equations are suggested

$$I_c = I_c' + I_c'' + I_c''' \tag{23}$$

$$I_c' = \frac{1}{\pi}\int_{\eta-\pi}^{\alpha} i_c d\omega t = -\frac{lH_d}{\pi N_c}\{[1 - \nu_d(1 + \cos\eta)](\alpha - \eta + \pi) - \nu_d(\sin\alpha + \sin\eta)\} \tag{32}$$

$$I_c'' = \frac{1}{\pi}\int_{\alpha}^{\zeta} i_c d\omega t = -\frac{xlH_d}{\pi N_c} \times [1 - \nu_s(1 + \cos\alpha + \cos\eta)](\zeta - \alpha) \tag{33}$$

$$I_c''' = \frac{1}{\pi}\int_{\zeta}^{\eta} i_c d\omega t = -\frac{xlH_d}{\pi N_c} \times [1/3 - \nu_s(1 + \cos\alpha + \cos\eta)](\eta - \zeta) \tag{34}$$

Oscillograms of gate and control currents are shown in Fig. 5 for the amplifier in doubler circuit, but with the entire load within a rectifying bridge. Here $R_c = 180$ ohms, $R_L = 800$ ohms, $L_L = 2.5$ henrys and $K = +0.20$.

A matter of concern is seen in the oscillations observed in all circuits and, in particular, in i_c upon the firing shock for this kind of load. The preceding equations were derived under the assumption that the effects of these oscillations average to zero in the integration process.

Fig. 5 presents also a comparison of calculated and measured transfer characteristics for the conditions described. Discontinuous flow occurs for $\alpha > 60$ degrees, as recognized with the procedures given in part I, sections 2 and 4, and reviewed in section 1 of part II. I_L is the sum of the forcing term I_{LS}' and of the relaxation term I_{LS}'' given by equations 24 and 25, part I, in cases of continuous flow and by twice the values of equations 12 and 13, part I, for cases of discontinuous flow. On the other hand, $I_G = I_{LS}' + I_{g1p}$ where I_{g1p} is the correction from prefiring currents in the gate winding, $I_{g1p} \cong 2lH_d[\alpha - (\eta - \pi)]/\pi N_g$. (In the case of discontinuous flow extinguishing at δ a similar correction can be made for I_L adding to it $2lH_d[\alpha - (\delta - \pi)]/\pi N_g$ to account for the fact that prefiring magnetizing currents flow in the load during the interval between extinction and firing.)

6. Self-Saturating Circuits with Additional Degenerative Feedback

The historical development of magnetic amplifier analyses, starting from the simple saturable reactor and successively adding feedback turns N_f, made it customary to say that the simple reactor has zero feedback and that the addition of, e.g., $N_f = 0.8\ N_g$ turns converts it into an amplifier with 80 per cent feedback. With this terminology conventional doublers and similar self-saturating circuits are amplifiers with 100-per-cent feedback.

The recognition that the simple satura-

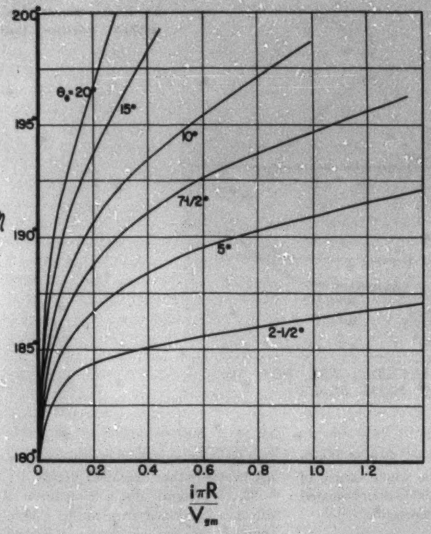
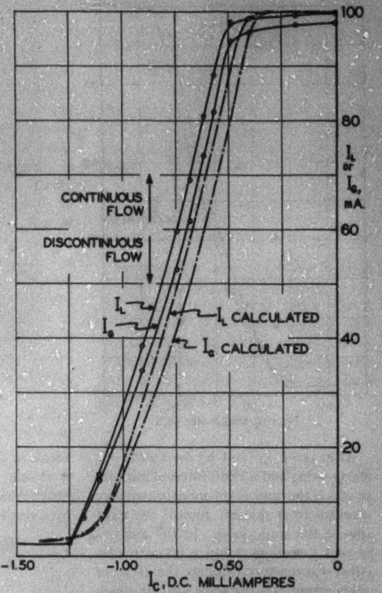


Fig. 4 (left). Extinction angle of gate currents, rectified loads with discharge paths



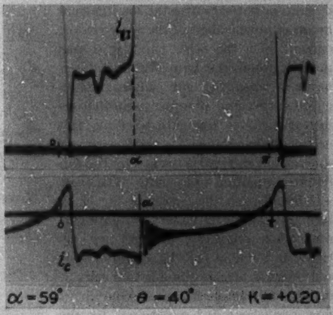


Fig. 5 (right). Rectified inductive load with discharge paths

ble reactor with sensitive core materials has unity ampere-turn gain later led to the following suggestion: it can be said more properly that the saturable reactor inherently has a 100 per cent feedback. The external feedback amplifier with $N_f = N_g$ (and thus the conventional self-saturating circuits) should be more properly considered as the basic amplifier. That is, an external feedback amplifier with $N_f = 0.8\ N_g$ has a degenerative feedback of 20 per cent. Likewise, if $N_f = 1.03\ N_g$, this amplifier has a regenerative feedback of 3 per cent.

Degenerative action is obtained in the self-saturating amplifiers by introducing mmf proportional to the rectified amplifier output current and acting on each core with a negative contribution to its total mmf lH. According to a suggestion by Krabbe[13] this can be obtained without need for other rectifiers, by adding N_d turns on each core; these turns are cross-connected in series with the gate winding on the other core with proper polarities. The degenerative action is seen from the recognition that the mmf equations 5 and 6 modify into

$$lH_{\mathrm{I}} = N_c i_c + N_g i_{g\mathrm{I}} - N_d i_{g\mathrm{II}} \tag{35}$$

$$lH_{\mathrm{II}} = N_c i_c + N_g i_{g\mathrm{II}} - N_d i_{g\mathrm{I}} \tag{36}$$

The operation of the amplifier in terms of total requirements of instantaneous mmf lH_{I} and lH_{II} remains unaffected (except perhaps for minor departures in the third interval if R_c is large). Considering the operation of core II during the half-cycle initiating at $\omega t = 0$ (or at $\omega t = \beta - \pi$ for the case of nonrectified, inductive-resistive loads) it appears now that the various mmf requirements of the core are satisfied for any given angle of firing if the original instantaneous control current i_c is varied by the addition of a term $i_c^* = i_{g\mathrm{I}} N_d / N_c$. That is, for a given α (and thus for a given output) the control current I_c is given by

$$I_c = I_c' + I_c'' + I_c''' + I_c^* \tag{37}$$

where, for each load situation, I_c', I_c'' and I_c''' are the contributions given to the total control current in the absence of degenerative turns by the three intervals considered in the previous sections, and $I_c^* = +i_{g\mathrm{I}} N_d / N_c$. The summation of $I_c' + I_c'' + I_c'''$ was seen to be a negative number; the additional requirements of degenerative action introduce a positive term. Thus the transfer characteristic expressing I_G where $I_G \neq I_L$ for the loads considered in section 5) is modified by a kind of "shearing" to the right, and positive values of control currents may well be needed over certain portions of the transfer characteristic.

More commonly, degenerative feedback is obtained by feeding the N_d turns on the two cores directly in series with a bridge rectifying the amplifier output current i_O substantially with the same results. (In either case the addition of the N_d turns modifies the voltage equations of the gate loops; this is a simple matter[6,14] which may be disregarded if $N_d << N_g$ and which, in any case, is not within the scope of this paper).

As the degenerative action is increased by making N_d larger and larger, the terms I_c', I_c'' and I_c''' become more and more insignificant with respect to I_c^* and no need may be felt for an accurate consideration of the minor loops described by the cores. Then, of course, coarser representations of the $\Phi = f(lH)$ relationship, (and in particular the single-valued representation with infinite permeability in the unsaturated region) are adequate, as used in section 1.

7. Rectifier Effects

Improper rectifier behavior is often responsible for some departures from the results of the previous analysis. Effects of rectifier capacitance are most elusive, and elementary analyses of equivalent linear circuits are inadequate for many obvious reasons. It is suspected that rectifier capacitance emphasizes unfavorably the high-frequency oscillations which are observed in all the various coupled circuits upon the firing shock, with asymmetrical current wave forms. Inadequate knowledge of both rectifier capacitance and magnetic nonlinearities along the bottom of the minor loop prevent the authors from attempting an analysis of these effects at this time.

On the other hand, effects of rectifier leakage are more readily accounted for. In fact, in cases of low control circuit resistances the reverse voltages which appear across the rectifier of the unsaturated core during saturation intervals are known in terms of output currents and circuit resistances, e.g., $v_{\gamma \text{II}} \cong -r_g i_{g\text{I}}$ for the doubler, or $v_{\gamma \text{II}} \cong -2v_g + r_g i_{g\text{I}}$ for the center-tap circuit, etc. If enough information is available to relate these voltages to the negative $i_{g\text{II}}$ of leakage, its mmf can be introduced in the expression of lH_{II}, thus modifying i_c during the intervals in which significant leakage occurs. Evidently this negative mmf has degenerative effects quite similar to those described in section 6. In fact, if a linearization in terms of a fictitious rectifier reverse resistance is admissible, rectifier leakage corresponds to a true degenerative feedback. The equivalent N_d and thus means for leakage compensations by the addition of regenerative turns have been calculated with simple results in terms of all pertinent parameters.[14]

8. Regenerative Feedback and Related Effects

Regenerative action is obtained in the self-saturating amplifiers by addition of N_r turns on each core; these turns are cross-connected in series with the gate winding N_g on the other core with proper polarities, namely, regenerative action results if the current $i_{g\text{I}}$ through these additional turns introduces a positive contribution to the mmf lH_{II} of the other core. The regenerative action is seen from the recognition that the mmf equations 5 and 6 modify into

$$lH_{\text{I}} = N_c i_c + N_g i_{g\text{I}} + N_r i_{g\text{II}} \tag{38}$$

$$lH_{\text{II}} = N_c i_c + N_g i_{g\text{II}} + N_r i_{g\text{I}} \tag{39}$$

It appears that the unchanged mmf requirements of the core II are satisfied for any given angle of firing if the original instantaneous control current i_c is varied by the addition of a term $i_c^{**} = -i_{g\text{I}} N_r / N_c$. That is, the control current I_c for given α is given by

$$I_c = I_c' + I_c'' + I_c''' + I_c^{**} \tag{40}$$

where I_c', I_c'' and I_c''' are calculated for the same load and angle of firing in the absence of regenerative action and $I_c^{**} = -I_{g\text{I}} N_r / N_c$. The summation $I_c' + I_c'' + I_c'''$ was seen to be a negative number; I_c^{**} is also negative. The regenerative action modifies the positive branch of the transfer characteristic expressing I_G with a kind of shearing to the left. The same effects are obtained if the N_r turns are fed directly by a bridge rectifying the amplifier output current i_G.

A somewhat similar "pseudo regenerative" action is recognized in center-tap amplifiers with inductive loads (or in external feedback amplifiers with inductive loads connected inside of the N_f loop) whenever no paths are offered for load relaxation by the additional provision of discharging rectifiers. In these cases, mentioned in section 6, part I, load relaxation takes place through the amplifier windings. In the case of the center-tap circuit the load current i_L'' of relaxation intervals finds two parallel paths through both gate windings, i.e., $i_L'' = i_{g\text{I}} + i_{g\text{II}}$.

While relaxation takes place, both cores are in the prefiring condition of swinging fluxes, their mmf are dictated by the ascending and descending sides of the dynamic major loop, i.e.,

$$lH_{\text{I}} = N_c i_c + N_g i_{g\text{I}} \cong +lH_d$$

$$lH_{\text{II}} = N_c i_c + N_g i_{g\text{II}} \cong -lH_d$$

whence, by summation

$$i_c \cong -\frac{N_g}{N_c} \frac{i_L''}{2}$$

In the case of continuous flow of load current these conditions prevail throughout the prefiring interval. Thus

$$I_c = -\frac{N_g}{2N_c} I_L'' + I_c'' + I_c'''$$

If the decay of relaxation current is dictated primarily by the load and is not greatly affected by the gate-winding resistance (and leakage inductance) and by resistances and voltages of control circuit, reflected into the gate circuit by the transformer action of both cores, then I_L'' is obtained from equation 25 part I. Equations 33 and 34, part II, give I_c'' and I_c'''.

If the flow is discontinuous, relaxation is limited to the interval from $\omega t = 0$ to $\omega t = \delta - \pi$. In this case

$$I_c = -\frac{N_g}{2N_c} I_L'' + \frac{1}{\pi N_c} \times \int_{\delta-\pi}^{\alpha} lH_{\text{II}} d\omega t + I_c'' + I_c'''$$

where lH_{II} is obtained from the descending branch of the major dynamic loop over the time interval from the end of relaxation to the next firing, and I_L'' is given by twice the value of equation 13, part I.

The discharge of the load current i_L'' through the gate windings has an action similar to the one noticed in the regenerative feedback insofar as it increases $|I_c|$ for given output. It differs from the true regenerative action as increases of $|i_c|$ are noticed in the prefiring rather than in the saturation interval.

True regenerative action by additional turns N_r as well as pseudo-regenerative action of gate-winding relaxation often affect the shape of the transfer characteristic so as to yield multivalued outputs for certain ranges of I_c. Calculation of such transfer characteristics by the procedures just outlined permits prediction of instabilities of this kind.

Conclusions

An analysis of load behavior upon firing has been presented in part I[1] for certain common types of loads. The corresponding signal requirement art investigated here, dealing specifically with two-core amplifiers of the low-control impedance types. Amplifiers with low ampere-turn gains have been considered first. The recognition of the inadequacies of the extension of simplified analyses to the more critical situations of amplifiers with very high ampere-turn gains has brought about a closer examination of the minor hysteresis loops described by the cores in the operation. Tentative equations have been presented for the control current requirements in the case of low control circuit resistance; the modifying influence of moderate forcing resistors in the control circuit has been examined and overlap and other related effects have been described.

The study was carried out first for purely resistive loads, as this seemed necessary at the present state of the art. Extensions to more general load situations have been made subsequently under use of the results of part I, and transfer characteristics have been obtained for the various cases.

Situations of less extreme gains (as resulting, e.g., from an additional degenerative feedback or from the use of inadequate rectifiers) have been discussed. Also a starting point for a further quantitative study of instabilities due to overregeneration of various kinds has been established.

A by-product of this approach is the tentative recognition of those core properties which seem to be most significant in the control of amplifiers with low control impedance; this may bring some clarification in the matter of pertinent core-testing procedures and in the study of the problem of drifts.

References

1. The Operation of Magnetic Amplifiers With Various Types of Load. Part I—Load Currents for Given Angle of Firing, **L. A. Finzi, R. R. Jackson.** *AIEE Transactions*, vol. 73, pt. I, July 1954, pp. 270–79.

2. On the Mechanics of Magnetic Amplifier Operation, **Robert A. Ramey.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1214–22.

3. The Magnetic Amplifier, **W. C. Johnson.** *Electrical Engineering*, vol. 72, no. 1, July 1953, pp. 583–88.

4. Saturable Reactors with Inductive D-C Load, **H. F. Storm.** *AIEE Transactions*, vol. 71, pt. I, Nov. 1952, pp. 335–43.

5. The Behaviour of a Series-Transductor Magnetic Amplifier with Directly-Connected or Rectifier-Fed Loads, **A. G. Milnes.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 99, pt. II, Feb. 1952, pp. 13–23.

6. Rules of Similitude for Magnetic Amplifier Systems, **L. A. Finzi, H. L. Durand.** *Proceedings*, National Electronics Conference, Chicago, Ill., vol. VII, 1951, pp. 498–514.

7. Theory of Magnetic Amplifiers with Square-Loop Core Materials, **H. F. Storm.** *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 629–40.

8. *Discussion* by **Harold W. Lord** of reference 7. *Ibid.*, p. 637.

9. *Discussion* by **R. A. Phillips** of reference 7. *Ibid.*, p. 639.

10. The Influence of Magnetic Amplifier Circuitry upon the Operating Hysteresis Loops, **Harold W. Lord.** *AIEE Transactions*, vol. 72, pt. I, 1953 (Jan. 1954 section), pp. 721–28.

11. *Technical Report No. 13.* Office of Naval Research, Washington, D. C., Contract *N7 ONR 30306* and *30308*, Project no. 075–272 and 275, June 1953.

12. An Instability of Self-Saturating Magnetic Amplifiers Using Rectangular Loop Core Materials, **S. B. Batdorf, W. N. Johnson.** *AIEE Transactions*, vol. 72, pt. I, July 1953, pp. 223–28.

13. The Transductor Amplifier (book), **Ulrik Krabbe.** Lindhska Boktryckeriet, Orebro, Sweden, 1947.

14. The Effective Feedback Ratio of Magnetic Amplifiers, **L. A. Finzi, G. F. Pittman, Jr., H. L. Durand.** *AIEE Transactions*, vol. 71, pt. I, April 1952, pp. 157–64.

No Discussion

Switching Functions on an n-Dimensional Cube

C. Y. LEE
NONMEMBER AIEE

THE analysis and synthesis of 2-terminal combinational networks from the standpoint of truth tables and Boolean algebra have been treated in great detail.[1,2] In this paper these networks will be considered from a somewhat different viewpoint, mainly a switching function will be considered as a set of vertexes on an n-dimensional cube (n-cube). Once a correspondence between a switching function and a set of vertexes on an n-cube is found, the concept of a "distance" may then be introduced. This concept proves to be of value in the problems of discovering symmetries and of classifying types of switching functions.

Distance and Symmetry on an n-Cube

Given n binary variables $(x_1, x_2 . . . x_n)$, there are 2^n "states" of these variables: $(x_1, x_2 . . . x_n)$, $(x_1', x_2 . . . x_n) . . . (x_1', x_2' . . . x_n')$. A binary n-digit point (vertex) on the n-cube corresponds to each of these states. Thus the state $(x_1', x_2\, x_3 . . . x_n)$ corresponds to the point (0, 0, 1. . .1). A switching function $f(x_1, x_2 . . . x_n)$ of n-variables takes on the value 1 for some of the 2^n states of these variables. (In this paper the convention is used that x is 0 or 1 according as the relay contact x is open or closed and f is 0 or 1 according as the input and output terminals are disconnected or connected.) Therefore the function f is represented by the set of points on the n-cube for which f takes on the value 1. Thus the switching function of three variables defined by the truth table, Table I, is represented by the three points on the 3-dimensional cube: (0, 0, 1), (0, 1, 0), (1, 0, 1).

Let $C(n)$ be the set of all binary n-digit points on the n-cube. A star (*) operation[3,4] is defined as follows. For every pair of points P_i and P_j in $C(n)$, $P_i{*}P_j=P_k$ where a digit of P_k is zero if the corresponding digits of P_i and P_j are the same, and a digit of P_k is 1 if the corresponding digits of P_i and P_j are different. Clearly, P_k is also a point of $C(n)$. Thus, if $P_i=(1011011)$ and $P_j=(0001101)$ (for $N=7$), then $P_i{*}P_j=(1010110)$.

Next, for every P_i in $C(n)$, $|P_i|$ is defined as the number of 1's in P_i. Thus, if $P_i=(1011011)$, then $|P_i|=5$. A distance can now be defined on the n-cube as follows.[5] For every pair of points P_i, P_j in $C(n)$, the distance between them is

$$d(P_i, P_j)=|P_i * P_j|$$

It can be shown that the distance axioms are satisfied.

A symmetry of the n-cube is any 1-to-1 transformation of the cube which leaves all pairwise distances invariant. It has been shown that there are $2^n(n!)$ symmetries of an n-cube. Thus, for $n=2$ there are exactly eight symmetries of the square (2-dimensional cube).[6]

Symmetric Switching Functions on an n-Cube

Before 2-terminal switching functions are discussed, the symmetric switching functions will be considered; as a subclass they are of considerable importance in practical application.[2,7] Geometrically speaking, whereas series-parallel networks correspond to subcubes[8] on an n-cube, symmetric switching functions correspond to "spheres" on the n-cube. Note particularly that both the subcubes and spheres possess a high degree of symmetry on an n-cube.

Given a set of n-variables $(x_1, x_2 . . . x_n)$, there are 2^n distinct transformations of this set of variables

$$(x_1, x_2 . . . x_n)\rightarrow(y_1, y_2 . . . y_n)$$

where each y_j is either x_j or x_j' for $j=1,2. . . n$. Let these transformations be called negations. A symmetric switching function of n-variables is then defined to be any switching function which remains unchanged under all permutations of some set of the variables $(y_1, y_2 . . . y_n)$. A better insight into symmetric switching functions can be acquired by introducing a class of elementary symmetric switching functions. For n-variables $(y_1, y_2 . . . y_n)$ where again each $y_j=x_j$ or x_j' for $j=1,2. . . n$, the elementary symmetric switching functions are denoted by $\sigma(n,i)$ where

$$\sigma(n,0)=y_1'\cdot y_2'\cdot y_3'\cdot \ldots \cdot y_n'$$
$$\sigma(n,1)=(y_1\cdot y_2'\cdot y_3'\cdot \ldots \cdot y_n')$$
$$+(y_1'\cdot y_2\cdot y_3'\cdot \ldots \cdot y_n')$$
$$+\ldots+(y_1'\cdot y_2'\cdot y_3'\cdot \ldots \cdot y_n)$$
$$\cdot$$
$$\cdot$$
$$\cdot$$
$$\sigma(n,n)=y_1\cdot y_2\cdot y_3\cdot \ldots \cdot y_n$$

The notation $\cdot$ is used for intersection and + for union. It is then seen that $\sigma(n,i)$ is the switching function which takes on the value 1 if i and only i of the n-variables $(y_1, y_2 . . . y_n)$ takes on the value 1. It can be shown that a switching function is symmetric if and only if it is a union of a subset of elementary symmetric switching functions.

The elementary symmetric switching functions are closely related with the Boolean operation $\oplus$ defined by

$$y_1 \oplus y_2 = y_1'\cdot y_2 + y_2'\cdot y_1$$

Instead of discussing the fundamental properties of symmetric and elementary symmetric switching functions individually, they will all be collected into one theorem.

THEOREM 1

1. A switching function is symmetric only if it is a union of a subset of elementary symmetric switching functions.

2. Denoting the complement of a switching function f by f', the following are true

(a) $[\sigma(n, i)]' = \sum_{k, k\neq i} \sigma(n, k)$

(b) If $f=\sigma(n,i_1)+\sigma(n,i_2)+\ldots+\sigma(n,i_k)$ (f is symmetric), then

$$f' = \sum_{k, k\neq i_1 \ldots i_k} \sigma(n, k)$$

Table I. Truth Table for Switching Function f

x_1	x_2	x_3	
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	0

Paper 54-207, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted February 19, 1954; made available for printing April 6, 1954.

C. Y. LEE is with the Bell Telephone Laboratories, Inc., New York, N. Y.

The author is indebted to H. N. Seckler, D. H. Evans, F. E. Hohn, S. H. Washburn, and W. Keister of the Bell Telephone Laboratories, Inc., for their many valuable suggestions and criticisms in the preparation of this paper.

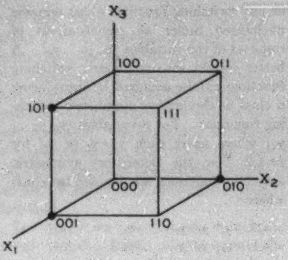


Fig. 1. Geometrical representation of switching function f

3. $y_1 \oplus y_2 + .\ .\ . \oplus y_n = \sigma(n,1)\sigma(n,3) + .\ .\ . + \sigma(n,n_0)$ where $n_0 = n$ if n is odd and $n_0 = n-1$ if n is even.

4. If we define $y_i^0 = y_i$, $y_i^1 = y_i'$, and $e_i = 0$ or 1, then $(y_1^{e_1} \oplus y_2^{e_2} \oplus .\ .\ . \oplus y_n^{e_n}) = y_1 \oplus y_2 \oplus .\ .\ . \oplus y_n$ if $\sum_{i=1}^{n} e_i$ is even, and $(y_1^{e_1} \oplus y_2^{e_2} \oplus .\ .\ . \oplus y_n^{e_n}) = (y_1 \oplus y_2 \oplus .\ .\ . \oplus y_n)'$ if $\sum_{i=1}^{n}$ is odd.

Now consider certain configurations on the n-cube which are the exact analogues of ordinary spheres.[5] Let P_0 be an arbitrary point of $C(n)$; then by a sphere of radius r about P_0 is meant the set of points P of $C(n)$ such that the distance $d(P, P_0)$ is r. In the case of ordinary spheres, the center of a sphere usually means a point which is equidistant from all points on the sphere. By this definition each sphere has two centers if infinity is considered as a point. In exactly the same way, every sphere on $C(n)$ has two centers, both of which are now finite. To distinguish them one will be called the near-center and the other the far-center.

As an illustration, consider the two spheres on the 4-dimensional cube consisting of the following two sets of points:

(a)	(0001)		(b) (1011)
	(0010)		(1101)
	(0100)	and	(0001)
	(1000)		(1110)
			(0010)
			(0100)

(a) is a sphere of near-center (0000) of radius 1 and of far-center (1111) of radius 3. The points (1000) and (0111) are again the near- and far-centers of (b) but the radius is 2 in both cases.

In general, if r is the radius of a sphere with respect to the near-center, $n-r$ is the radius with respect to the far-center where $r \leqq n-r$.

If a 2-terminal switching function is considered as a set of points on an n-cube, then an elementary symmetric switching function $\sigma(n,i)$ corresponds exactly to a sphere on $C(n)$ with a radius of the smaller of i and $(n-i)$ with respect to its near-center. (For example, the elementary symmetric function $\sigma(5,3)$ corresponds to a sphere on $C(n)$ with radius 2 with respect to its near-center.) The following theorem results.

THEOREM 2

Every elementary symmetric switching function $\sigma(n,i)$ corresponds exactly to a sphere on the n-cube (Cn) and conversely. If P_0 is the near-center, the radius of the sphere with respect to P_0 is the smaller of the two numbers i and $n-i$. If Q_0 is the far-center, Q_0 is the "complement" of P_0 [e.g., if $P_0 = (0111)$, then $Q_0 = (1000)$].

The problem of synthesizing symmetric switching networks becomes now a problem of recognizing spheres on an n-cube. A method of recognizing such spheres by using a map has been described by Caldwell.[7]

Types of Switching Functions and the n-Cube

By a symmetry (sometimes referred to as an input transformation[8]) of an n-cube is meant a 1-to-1 transformation of the cube which leaves all pairwise distances on the cube invariant. Let G be the set of all symmetries of an n-cube and let γ be any element of G. If f is any switching function of n-variables, then γ, operating on the cube, induces a new switching function g of n-variables. This operation is denoted by $\gamma(f) = g$.

For example, let f be the switching function of three variables previously given and let γ be the symmetry operation

$y_1 = x_3$, $y_2 = x_2'$, $y_3 = x_1$

$\gamma(f)$ is then the switching function g given by the truth table shown in Table II. A geometric picture is shown in Figs. 1 and 2.

Two switching functions f and g are said to be of the same type if there is some symmetry γ in G such that $\gamma(f) = g$.

Table II. Truth Table for Switching Function g

y_1	y_2	y_3	g
0	1	0	0
1	1	0	1
0	0	0	1
1	0	0	0
0	1	1	0
1	1	1	1
0	0	1	0
1	0	1	0

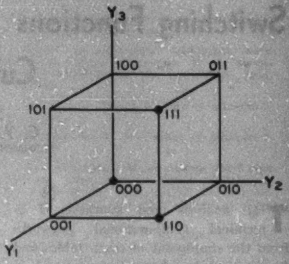


Fig. 2. Geometrical representation of switching function $g = \gamma(f)$

Given two functions f and g, it is important to know whether or not they are of the same type, because the synthesis problem for f is equivalent to the synthesis problem for g if f and g are of the same type. Thus, although there are a total of 65,536 switching functions of four variables, the synthesis problem is reduced to the synthesis of each of the 402 types of switching functions of four variables.

Mathematically, the set G forms a non-Abelian group and is called the group of symmetries of the n-cube. By using the theory of groups, the number of types of switching functions of n-variables has recently been evaluated.[10,11]

There are still many open problems in this connection. One problem is, given a switching function f, to find if there is any nonidentity symmetry γ in G such that $\gamma(f) = f$ and, if so, how many such symmetries there are.[9] A more important problem is how the theory discussed here can be applied to sequential relay networks. It is hoped that solutions to these and other similar problems will not be too distant in their appearance since even a partial solution to these problems may lead to better and more economical design of practical circuits.

References

1. THE DESIGN OF SWITCHING CIRCUITS (book), **William Keister, A. E. Ritchie, S. H. Washburn.** D. Van Nostrand Company, New York, N. Y., 1951.

2. SYNTHESIS OF TWO-TERMINAL SWITCHING CIRCUITS, **C. E. Shannon.** *Bell System Technical Journal*, New York, N. Y., vol. 28, 1949, pp. 59–98.

3. BINARY CODES WITH SPECIFIED MINIMUM DISTANCE (book), **M. Plotkin.** Moore School of Engineering, University of Pennsylvania, Philadelphia, Pa., 1952.

4. ON THE TYPES OF COMPOUND STATEMENT INVOLVING 4 CLASSES, **W. K. Clifford.** *Mathematical Papers*, London, England, 1882.

5. ERROR DETECTING AND ERROR CORRECTING CODES, **R. W. Hamming.** *Bell System Technical Journal*, New York, N. Y., vol. 29, 1950, pp. 147–60.

6. SURVEY OF MODERN ALGEBRA (book), G. Birkhoff, S. MacLane. The Macmillan Company, New York, N. Y., 1953.

7. THE RECOGNITION AND IDENTIFICATION OF SYMMETRIC SWITCHING FUNCTIONS, S. H. Caldwell. *AIEE Transactions*, vol. 73, pt. I, May 1954, pp. 142–47.

8. THE MAP METHOD FOR SYNTHESIS OF COMBINATIONAL LOGIC CIRCUITS, M. Karnaugh. *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 593–99.

9. SYNTHESIS OF ELECTRONIC COMPUTING AND CONTROL CIRCUITS (book), Staff of the Harvard Computation Laboratory. Harvard University Press, Cambridge, Mass., 1951.

10. ON THE NUMBER OF SYMMETRY TYPES OF BOOLEAN FUNCTION OF N-VARIABLES, D. Slepian. *Canadian Journal of Mathematics*, Toronto, Ont. Canada, vol. 5, 1953, pp. 185–92.

11. THE NUMBER OF STRUCTURES OF FINITE RELATIONS, R. L. Davis. *Proceedings*, American Mathematical Society, New York, vol. 4, June 1953, pp. 486–95.

◆

No Discussion

Varistor Modulators for Carrier Systems

R. S. CARUTHERS
MEMBER AIEE

SINCE the advent of the varistor modulator in carrier communication equipment about 20 years ago, much has been written on the theoretical nature of the modulation processes involved. Very little has been written with regard to the practical problems in modulator use encountered by the designer. This paper will stress the practical rather than the theoretical.

Early varistor modulators were all of a copper-oxide type;[1] in fact, little else was available other than the cat-whisker crystals used in radio sets. More recently newer systems have almost abandoned the use of copper oxide in favor of point-contact varistor diodes of the germanium and silicon types. These later units show much improved frequency characteristics, better stability, and lower cost. Their construction, too, is more compatible with present-day miniaturization approaches.

Modulators in speech communication equipment take many forms. In channel modulators and demodulators, they serve not only to modulate the speech band with a single carrier frequency but also as a limiter. In group modulators, a considerable number of channels are modulated on a carrier for proper disposition of the channels in the frequency spectrum of the transmission medium involved. In carrier telephone repeaters for open-wire and cable systems, a wide use has been found recently in providing a frequency-frogging modulator arrangement whereby the group spectrum is changed alternately between low- and high-group bands at successive repeater points.[1] Varistor modulators find other uses as keying devices either in the signaling branches for on-off operation in single-frequency signaling or in telegraph. In frequency-shift signaling operation an oscillator may be keyed back and forth between two frequency positions by varistor operation of a frequency-determining element in the oscillator circuit. In this paper the operation of channel and group modulator circuits will be stressed, as well as their influences on the operation of signaling and regulating circuits and on the design of associated filter equipment.

Channel Modulators

Design of channel modulators is concerned with:

1. Choice of the frequency allocation.
2. Selection of a plan of modulation.
3. Choice of a proper modulator configuration.
4. Evaluation of a procedure for supplying carrier to both modulator and demodulator.
5. Obtaining proper loading and modulation effects in the modulator.
6. Working out band filter and voice filter arrangements for association with the modulator.
7. Obtaining proper crosstalk either from channel to channel through proper filter design or through common carrier supply arrangements.
8. Properly associating the whole into a channel of adequate frequency characteristic and distortion-free performance.

In addition, where signaling channels are involved, signaling must be properly integrated with the speech modulator so that the interacting effects may be controlled adequately.

Fifteen years ago a principal concern was that of obtaining low loss in modulators; this has since become of negligible importance. Modulator circuits, when combined with group modulating equipment which necessarily works at low levels, are able to make liberal use of pads for level co-ordination, and modulator loss is not important. In demodulators, because of the high output levels required at terminals, it has never proved feasible to operate without a voice amplifier after the demodulator unit. This gain makes the loss of the demodulator relatively unimportant.

Either shunt, series, or double-balanced (ring-type) modulators are employed. Use of a minimum number of transformers, obtaining a proper overload characteristic for adequate limiting, and freedom in use of grounded circuits are major factors in selecting the proper form of modulator. Because it is not economical to design associated voice-frequency filters and channel band filters (coil-capacitor types are referred to) on any other than an unbalanced basis, transformers are usually required either beyond the filters or within the modulator structure.

In Fig. 1 channel modulators of both shunt and ring type are shown. In both cases transformers are needed toward the input voice circuit because of requirements for balanced operation in connection with connected 4-wire terminating sets. The band filter end of the modulator is unbalanced to ground in each case to cheapen the band filter design. In the case of the ring modulator this necessarily imposes a transformer requirement in the carrier supply branch. A transformer also is required in the carrier supply branch because of the shunt modulator. With grounded operation in the speech branches of the circuit, crosstalk results from channel to channel through the carrier supply branch unless a transformer winding is imposed in each modulator branch separately fed from the carrier supply.

A factor of importance in the design of channel modulators is the degree of overloading or clipping presented to the high speech peaks in the signal. The objective for such a limiting device is that strong signals from an individual channel will not seriously load subsequent common amplifiers and modulators for groups

---

Paper 54-218, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 22, 1954; made available for printing April 13, 1954.

R. S. CARUTHERS is with the Lenkurt Electric Company, Inc., San Carlos, Calif.

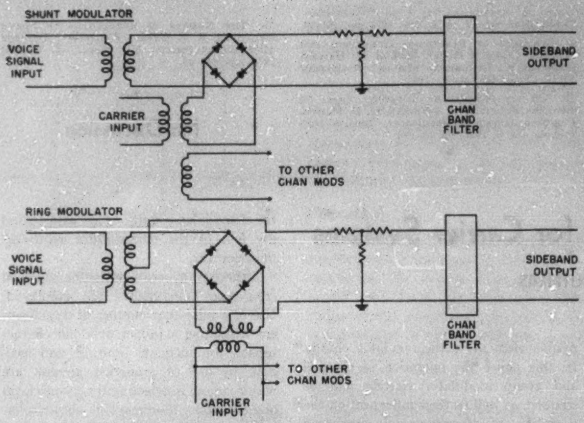


Fig. 1. Schematic diagrams of shunt and ring-type modulators

of channels. This prevents any one channel from rendering the multichannel system inoperative. Strong signals from a channel may come from:

1. High speech peaks associated with the strongest of talker volumes.
2. High signaling levels experienced with ringdown signaling.
3. Switch hook or relay clicks.
4. Accidental subjecting of the modulators to high-level single-frequency signals during test and maintenance.

For many years the degree of limiting put in a speech channel in carrier telephone communication equipment was set by the overload characteristic of three *44-A1* Western Electric repeaters operated in tandem. Such a load curve is shown in Fig. 2. It was reasoned that long voice-repeatered circuits operated satisfactorily with this degree of limiting. Limiting of this magnitude was such that after the first three *44-A1* repeaters in tandem any number of subsequent repeaters changed the gain output curve by an inappreciable amount. The only change in this limiting curve in present design objectives is that talker volumes have decreased over the last 15 to 20 years to an extent that output can be limited more than is shown by this curve. At the time of the initial use of this curve in setting modulator characteristics, average talker volumes in telephone circuits were considered to be at least 6 decibels (db) greater than they are today. Generally, however, the maximum output of channel modulators is limited only about 3 db more than in design procedures of 15 to 20 years ago. In military equipment, where the speech levels at the input of modulator equipment are not subject so such wide variations as on associated trunk circuits, limiting has sometimes been made 6 db more severe than for three *44-A1* repeaters in tandem. Part of the hesitance in putting more severe limiting in telephone practice is caused by the fear that lack of maintenance will allow input levels to channel modulator equipment to exceed appreciably the normal values used in design. In Fig. 2 an overload curve for a shunt modulator is shown closely approximating the shape of three *44-A1*'s.

In Fig. 3 overload curves are shown for ring modulators for various amounts of carrier voltage and power. The degree of limiting is determined by the carrier voltage and power supplied. It should be noted that ring modulators do not overload as abruptly as the shunt modulator or the three *44-A1* objective. The more gradual overloading is caused by higher impedance in the carrier supply branch than in shunt modulators.

Choice of Channel Modulator Frequency Allocation and Modulating Plan

Basically, frequency allocations for channel modulators revolve about the types of filters which will be used. Use of crystal filters influences the design toward choice of carrier frequencies in the range between 100 and 200 kc. Use of coil-capacitor filters influences the design to employ the lowest possible frequencies to obtain adequate *Q*'s and good stability of elements.

Fig. 4 shows a modulating plan setting up 12 channels in the 40-to-88-kc band. Channels are made up in groups of four, 4-kc spaced, and occupying a spectrum from 8 to 24 kc. Use of such a low-fre-

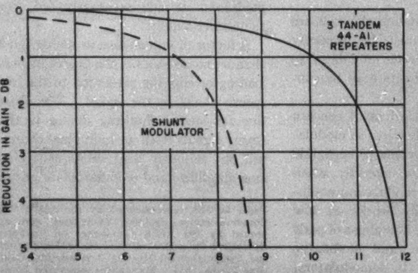


Fig. 2. Overload characteristics for three Western Electric type-44-A1 repeaters in tandem and for a shunt-type modulator

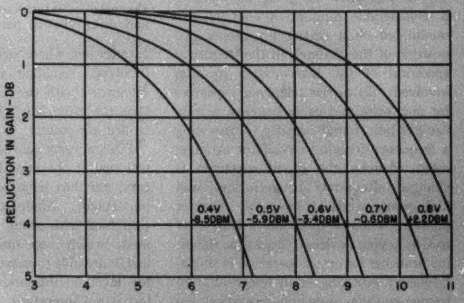


Fig. 3. Overload characteristics of a ring modulator for various amounts of carrier voltage and power

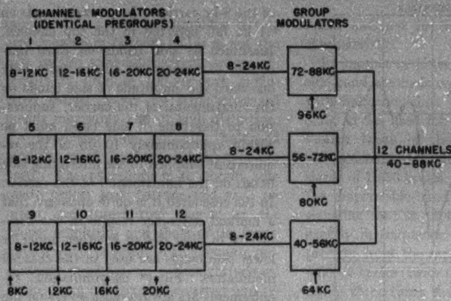


Fig. 4. A modulation plan for generating 12 channels at 40 to 88 kc through use of three 4-channel pregroups

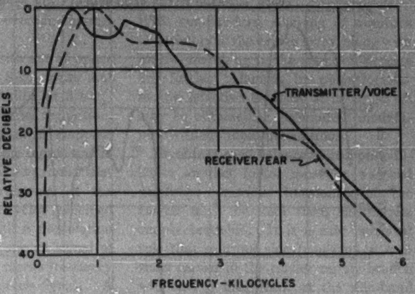


Fig. 5. Transmitter-voice and receiver-ear characteristics for the Western Electric F1A telephone set

quency band as this with carriers at 8, 12, 16, or 20 kc inevitably results in production of side band on harmonics of these carriers not much removed in frequency from the useful bands. In fact, side bands on harmonics of the carriers of the lower channels overlap the side bands of the upper channels. This is not the case where the primary modulating bands are chosen at the high frequencies used with crystal filters. A sufficiently high channel modulating band frequently results in utilization of a common filter to a group of channels, suppressing side bands of harmonics of any carrier within the group of channels. To permit use of a common low-pass filter to suppress all such harmonic side bands, it is necessary only that the side band on the second harmonic of the carrier of the lowest channel involved fall above the side band of the carrier on the highest frequency channel. This is particularly useful in double-side-band compandored systems where, as a result, no transmitting channel band filter is needed.

Channel Band Filter Requirements in Relation to Channel Modulator Arrangements

In the design of channel band filters for association with modulator circuits, use is made of the so-called transmitter-voice and receiver-ear frequency characteristics. In most instances both side bands of the transmitting channel frequencies are displaced with respect to the adjacent receiving channel band. As a result of this staggering, a disturbance advantage results because the highest energy content of the voice in the transmitting channel is not heard at the most sensitive frequency spot of the listener ear in the receiving channel. A common procedure for many years has been to break the energy content in the transmitter-voice characteristic into 100-cycle segments with a measured amount of total energy assigned to each. Likewise, the relative sensitivity in the receiver-ear characteristic is compared in 100-cycle segments across the band of the receiving channel.

Thus, in two adjacent channels 4 kc spaced and operating on lower side bands, a 1,000-cycle speech signal modulates the carrier to produce upper and lower 1,000-cycle side bands. In the adjacent receiving channel these two side bands in the interfering channel are heard respectively at either 3 or 5 kc. The relative interfering effect into the disturbed channel is determined by a comparison of the sum of the original 1,000-cycle signal energy and the 3-kc receiver-ear weighting in the receiving channel with the sum of the transmitter-voice and the receiver-ear energy weighting of the 1,000-cycle tone in the disturbing channel.

In Fig. 5 transmitter-voice and receiver-ear characteristics are shown for the Western Electric type-*F1A* telephone set. The receiver-ear characteristic shown is quite similar to the weighting curve commonly used with noise meters of the Western Electric *2B* type. The total speech power output of the transmitter, expressed as the sum of the 100-cycle transmitter-voice segments, is 10 db greater than the power of the maximum segment at 600 cycles. The average talker volume at 2-wire input to carrier channels in telephone plant is about −16 volume units (approximately −16 db below 1 milliwatt) with a standard deviation of about 5.5 db. The maximum talker volume is assumed to be 0 volume units and the minimum about −31 volume units.

In Fig. 6 transmitter-voice and receiver-ear characteristics have been used to compute the filter discrimination needed in channel band filters with carriers 4 kc apart to result in 55-db crosstalk interference between these channels. Each interfering 100-cycle segment was assumed to contribute equally to the interference. A total of 40 100-cycle segments was assumed in summing the power contribution in the interference

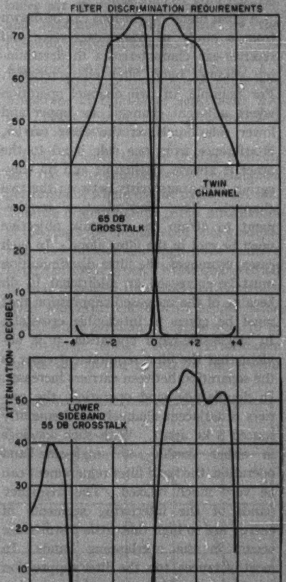


Fig. 6. Filter discrimination requirements for channel band filters

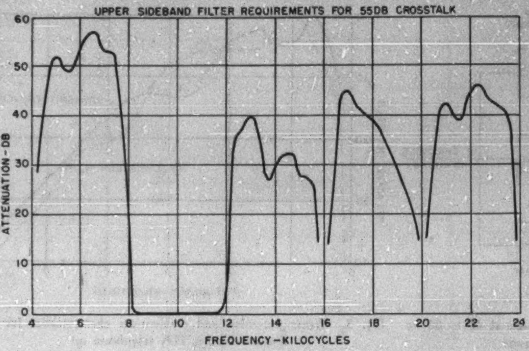


Fig. 7. Filter discrimination requirement for a channel band filter using upper side-band channel

band. It should be noted that, even though the crosstalk requirement is 55 db between such channels, in only a few of the 100-cycle segments is the filter requirement as great as 55 db. In the majority of cases the staggering advantage due to displaced frequencies permits relaxed filter requirements.

A great many conclusions can be drawn from a study of transmitter-voice and receiver-ear characteristics in determining channel band filter characteristics. For example, in twin-channel operation where adjacent channels are upper and lower side bands on the same carrier, interference from one side band to the other is always intelligible and no staggering advantage exists between the two channels. The discrimination requirement to attain the crosstalk objective must be met in the filter alone. In such cases, moreover, the filter discrimination must be increased an additional 10 db because of the increased suppression that must be given to intelligible crosstalk. In single-side-band operation, it is obvious that the filter requirement eases as the separation between carriers increases. In double-side-band operation, the carriers of adjacent channels are frequently spaced 8 kc apart. With such spacings in either double- or single-side-band operation, the band filter requirement can be very much relaxed. The frequency bands of the interfering segments of energy are so high that little disturbance occurs in the overlapping bands. In most instances, too, the filter suppression required may be put in the voice-frequency band with low-pass filters either in the transmission circuit or in the receiving circuit. This virtually eliminates any need for band filters in double-side-band operation. It is advantageous in such cases to minimize even further the amount of filtering required, through resort to compandors or through use of phase modulator methods. In the case of compandors as much as 20 to 25 db of additional suppression is obtained, and in the case of phase modulators the major part of the unwanted side band is easily reduced about 30 db.

In Fig. 7 filter requirements which will result in no greater than 55-db interference into adjacent channels in an 8-to-24-kc 4-channel group have been computed for an 8-to-12-kc band filter. Upper side-band arrangements are used with carriers at 8, 12, 16, and 20 kc. Any direct interference from the lower channel into the next higher channel requires speech band frequencies above 4 kc modulating on the 8-kc carrier. However, lower side bands on the 8-kc carrier are heard with much greater intensity on the next lower frequency channel after group modulation (see Fig. 4), 3-kc side bands, for example, being heard at 1,000 cycles. In the higher frequency channels direct interfering products result from the 8 kc carrier channel modulator through side bands on the second harmonic of 8 kc, lower side-band frequencies falling in the band from 12 to 16 kc. Upper side bands on this second harmonic fall in the band from 16 to 20 kc and are directly intelligible.

The shunt modulator side bands on the second harmonic of the carrier are seldom more than a few db below the side bands on the fundamental of the carrier. In a double-balanced or ring modulator an additional 25 db of suppression results for such side bands, and filter requirements can be reduced correspondingly. A further side-band interfering pattern occurs in this case on the third harmonic of the 8-kc carrier. Lower side bands on 24 kc fall inverted in the top channel. Inasmuch as side bands on third harmonics of carriers are theoretically only one-third the amplitude of side bands on the fundamental of the carrier, suppression to this type of side band must be within approximately 10 db of the requirement supplied to the interfering side bands on the fundamental of the carrier. In the case cited it is quite apparent that a marked difference results in the complexity of filter design whether upper or lower side bands are used in the channel modulators. Filter discrimination requirements are always most severe on the unwanted side-band frequency portion of the band. Beyond the band in the frequency portion of the wanted side band, interfering effects are small. Also, interfering effects with side bands on harmonics of the carrier are inclined to be large. It is quite apparent, then, that where extensive filter discrimination is required it is best to operate with lower side-band arrangements so that unwanted side-band discriminations and discriminations to side bands on harmonics of the carrier fall in the same frequency position or order. This is the case only with lower side-band operation.

Another filter requirement always present in operation of single-side-band channels is the discrimination needed for the unwanted side bands to prevent wobble in the transmission characteristic. It is seldom possible in suppressed carrier single-side-band operations to hold the carriers at the two ends in exact synchronization. Frequency deviations up to 10 cycles are normally tolerated in over-all speech channel bands, and up to as much as 2 cycles in program channels. It has been determined by ear tests that a frequency difference of about 4 cycles between the carriers at the two ends is most inclined to result in a disturbing effect. If the lower frequencies in the speech band are not adequately suppressed in the unwanted side-band portion of the filter, a beat will occur in these frequencies at the receiving end. A minimum suppression to unwanted side bands in the combined transmitting and receiving circuit should be at least 20 db and preferably from 30 to 40 db.

In the operation of double-side-band systems with the carrier transmitted, it is a simple matter to approach 100-per-cent modulation with very low distortion through the use of full-wave linear demodulator circuits. As theoretically predicted, distortion stays low until 100-per-cent modulation is exceeded in the demodulator, where the modulation breaks rapidly as in a feedback amplifier circuit near maximum output. This is not the case in the modulator circuit where attempts to produce 100-per-cent modulated carrier waves by straightforward means invariably end in high distortion. To overcome this difficulty balanced modulators of the ring or shunt type are used. The procedure is to obtain adequate carrier leak balance and then reinject the transmitted carrier at the proper amplitude and phase in the transmitted wave, to result in an adequate percentage modulation. Straightforward means have reinjected the carrier and resulted in poor control of the carrier phase stability from a time and temperature standpoint, so that the carrier phase is displaced with respect to the side-band phase.

It was found that the introduction of a d-c signal in the speech branch with an amplitude equal to the maximum speech peak resulted in production of two zero-frequency side bands which were identical with the desired transmitted carrier in frequency and phase. The end result was a modulator circuit with approximately 100-per-cent modulation on maximum speech peaks with the carrier wave under good control of phase.

Some unique experiences have been encountered in modulators where a tone-operated frequency-shift signaling channel is also used for automatic transmission regulation of the over-all speech channel. In this case the level of the signaling tone was used as an over-all means of adjusting receiving channel gain to offset line or transmitting and receiving terminal variations. Overloading from speech or test signals caused corresponding gain reductions in the level of the regulating signal tone. The channel regulator automatically would tend to compensate for such gain reduction and thus offset any overloading caused by the speech signal. Thus, no measure of the overload characteristic of the channel could be made on a single-frequency basis. Fig. 8 shows the overload of the channel modulator plus the reduction in signal level passing through the modulator. The reduction of gain at the signaling or regulating frequency is even greater than that at the speech signal. The end result, then, was that the gain of the over-all channel actually increased as the speech signal was increased. Even though the channel modulator at the transmitting end was being used as a limiter to clip the maximum peaks of the speech signals, actually on a single-frequency basis this effect could be measured only at the transmitting terminal and not in the over-all channel characteristic.

In another instance, with use of upper side-band channels on a 8-kc carrier and with frequency-shift signaling at 3,400/-3,550 cycles used for channel regulation, violent beats in the channel response were encountered with an input speech signal of 1,550 cycles. It was found that the upper side-band frequency of the signaling tone at 11,400 cycles coincided exactly in frequency with a modulation product from the input speech signal of 4,600/3 cycles (twice the carrier frequency minus three times the speech signal). A shunt modulator was used in which this modulation product was not balanced from the output. Even with such a product 26 db below signaling side band, a beat of 1 db was experienced which was further doubled through the use of a compandor in the channel. Use of a ring-type channel modulator, with its added 25 db of balance to such products, greatly reduces such troubles. Much of the time spent in development of modulator circuits is expended on some such unpredicted occurrences.

Group Modulation Design Considerations

A primary consideration in the design of group modulator circuits is to choose the frequency allocation so that the input frequency band will not overlap the output band, either through direct transmission or through modulation on any harmonic of the carrier. Invariably designers require the output band to be displaced with respect to the input band because of inability to hold more than about 25 db of balance in the modulator over the total input signal band. Likewise, lower side-band group modulators are generally used because they automatically result in nonoverlapping bands through modulation on harmonics of the carrier. With compandors, stacking of numbers of varistors in the bridge arms, along with careful matching in the selection process and further balancing by potentiometer means, effective balances of 60 or 70 db may be obtained in germanium modulators without these precautions.

To obtain good balance in group modulators another factor must be watched carefully: spurious harmonics of the carrier must be kept from the carrier supply branches. It is a safe universal rule that spurious side bands on a spurious carrier in the carrier supply branch will have a magnitude at the side-band output branch of the modulator reduced below the magnitude of the wanted side-band output by the same number of db that the spurious carrier is below the wanted carrier. This can be proved in simple mathematical terms. An output side band on the third harmonic of the carrier ($3f_c \pm f_v$) is one-third of the amplitude of the wanted side band ($f_c \pm f_v$). Side bands on third-order combinations of the carrier in the carrier supply branch are $2f_c \pm f_s \pm f_v$ and ($2 \times dc$ or zero frequency $\pm f_s \pm f_v) = f_s \pm f_v$. Their amplitudes are related to the third harmonic and fundamental side bands by the following approximations:

Amplitude of fundamental side band: k
Amplitude of third harmonic side band: $k/3$
Amplitude of third-order side band: $k \times C_s/C$ where C is the carrier amplitude and C_s the spurious carrier amplitude

Until recently the use of ring modulators was quite satisfactory in modulation performance where copper oxide was used. The use of germanium group modulators of the ring type encountered poor modulation performance and high noise relative to copper-oxide types. Noise nearly equal to that expected from thermal contributions was experienced with copper-oxide types, while germanium types were 10 to 15 db more noisy. Noise performance curves are shown in Fig. 9 for several

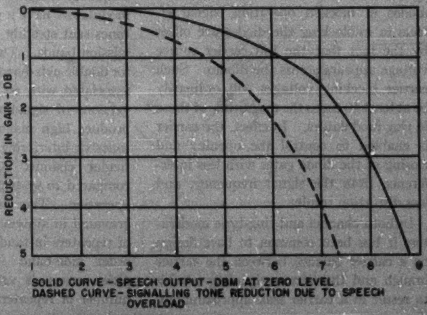


Fig. 8. Overloading effects on the regulating signal in the channel modulator caused by high-level speech signals

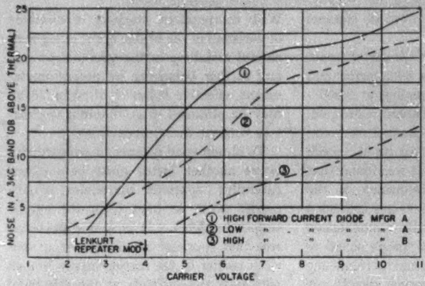


Fig. 9 (above). Comparison of modulator noise performance using germanium diodes

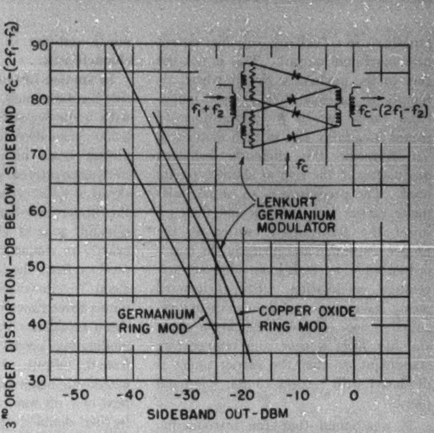


Fig. 10 (right). Comparison of modulators from a distortion standpoint

types of modulator units from germanium suppliers. Also shown is noise performance with the Appert modulator circuit.[2] In Fig. 10 modulation performance with copper-oxide and germanium ring modulators is shown under quite comparable carrier supply conditions. Neither noise nor modulation performance could be made satisfactory in germanium modulators at the signal levels found satisfactory in copper-oxide units. Germanium modulator arrangements of the type shown in Fig. 10, with high resistances in the carrier supply branches, result in much improved modulation performance and lower noise. The lower noise results from the fact that the carrier does not have to be raised to such high levels for adequate modulation performance. A modulator arrangement of this type gains its low-distortion advantage from the fact that the current flow from the carrier is maintained at nearly constant value due to the high series resistors, and the varistor resistances stay at values set by the carrier regardless of the loading effects of the speech signals. Also, during those portions of the carrier cycle when the carrier maintains the diodes in blocked operation, signal effects in unblocking the diodes are offset by the fact that the full carrier source voltage appears across the diodes. Such carrier blocking voltages will ordinarily be at least 10 times the magnitude of those in ring modulators. In effect, the carrier is enabled to control the opening and closing of the diode gates with less interference from the signal frequency, and less distortion results.

In both channel and ring-type modulators it has been common to base design on a carrier balance between the carrier branch and the output signaling branch to result in a carrier leak approximately 40 db below the input carrier power. At least 12 db of this balance figure are inherent in the bridge structure used; hence a 40-db balance does not reflect a closer control of diode characteristics than about 5 per cent. With the Appert modulator circuit and the same degree of varistor selection, carrier balances are improved 10 to 15 db. This results from the greater uniformity of diode impedance and match inherent with carrier current held constant in the bridge arms.

Exceedingly good frequency characteristics are obtained with germanium modulators and associated filters and amplifiers. Fig. 11 shows a characteristic of a frequency-converting repeater for cable circuits with a 24-channel input band from 40 to 140 kc and an output band from 164 to 264 kc. The transmission variation of earlier copper-oxide types has been largely eliminated. No equalization was necessary other than that required for the corners of the filters used.

Some of the more recent systems with frequency-frogging repeaters using group modulators have a great many high-level tones sent steadily along with the transmission bands. These are either carriers for double-side-band transmission or tones associated with on-off or frequency-shift versions of signaling. The tendency to produce high magnitudes of third-order tones or interfering side bands in a particular channel is greatly aggravated compared to systems with low-level tone operation. This problem is further aggravated in systems using large numbers of repeaters in tandem where such third-order side-band interference adds in phase or on a voltage basis with the number of repeaters. As an example, a system with 10 transmitted tones in one repeater would produce side-band interference amplitudes from a particular channel into another approximately 10 to 20 db greater than with only one tone present, and 10 such repeaters would further increase this modulation interference another 20 db. Both compandors and carefully controlled levels through the modulators and associated amplifiers have been found necessary in systems using such repeaters to keep the design under reasonable control.

Diode Selection for Modulator Use

One hesitates to describe a particular type of diode selection for a particular modulator use because of the variation in the source of supply and in the use of the particular modulators. In most instances diode selection for modulators is made by measuring the current at a particular voltage and binning diodes into groups with standard selection ranges. A balanced modulator is made by using four diodes from a particular bin. A distribution of diodes at a particular time at the Lenkurt Electric Company, Inc., is shown in Fig. 12. The diodes in the most concentrated area are chosen and used with other selections at a second current point for variolossers in compandor circuits. Such 2-point selection has resulted in very good uniformity of characteristics for this application. The great difficulty with this procedure is that manufacturers tend to remove the most uniform of the diodes to sell at premium prices in special customer applications. This spreads the distribution of diodes abnormally so that the eventual user no longer finds a high concentration

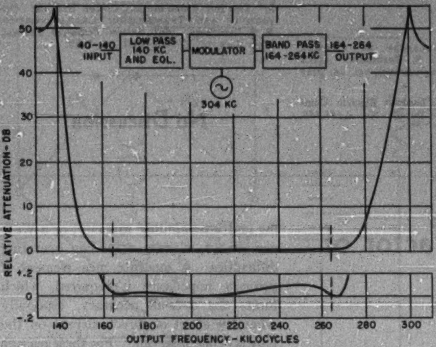


Fig. 11. Transmission characteristic of a frequency converting repeater for a 24-channel cable system

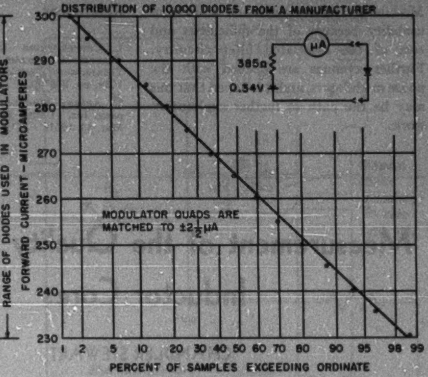


Fig. 12. Statistical distribution of diodes for modulator use

of good performance diodes near the average characteristic. It appears that the only sure way of getting a high quantity of well-matched diodes in the most desirable operating areas is to have the manufacturers select such diodes at the source of manufacture.

Flow of Modulation Products

In shunt, series, or ring-type modulators, a knowledge of the paths for flow of modulation products is an essential part of the design. Fig. 13 shows such modulation flow in a ring modulator. The basis for this direction of flow of modulation products is simply derived. The direction of input voice signal and carrier branch signal is assumed at a particular instant of time, as shown by arrows. The current-voltage relationship in each diode of the bridge is represented by a power series as shown in the figure. The resulting direction of flow of modulation product currents in each arm of the bridge is shown by the arrows. It is a simple matter to see, from the direction of current flow for each class of modulation product, into which branches of the array the various classes fall.

The power series approach and derivation of branch distribution of modulation products also make it simple to determine the effectiveness of balancing potentiometers to balance classes of modulation products from the four important modulator branches, namely the input and output signal branches, the carrier supply branch, and the ring. These potentiometers are indicated in the circuit diagram of Fig. 13. The potentiometer at *A* balances the carrier and all products of the $n_oc \pm n_ev$ class from the output signal branch. Likewise, it balances all products of the $n_oc \pm n_ov$ class from the carrier branch. It has no effect on balancing the input signals from the ouput branch. Likewise, the potentiometer at *B* balances the output branch signal of the $n_oc \pm n_ov$ class from the input signal branch, and vice versa. A potentiometer at *D* balances carrier branch classes of products from the input signal branch, and vice versa.

Conclusion

It has not been possible to describe in a paper of this length the modulator problems associated with diverse devices such as compandors, varistor keyers, and various signalling and telegraph arrangements. In these areas, too, there are many design problems of a practical nature that it would be of interest to describe.

Another area where a great deal of unexplored circuitry lies is in the realm of phase modulators or modulators where varying magnitudes of phase shift are put into the input signal branches and the carrier branches, to end in elimination of the unwanted side band by other than filter means or otherwise to obtain unusual operating characteristics. Many varieties of this type of modulator circuitry could be used to gain system advantage.

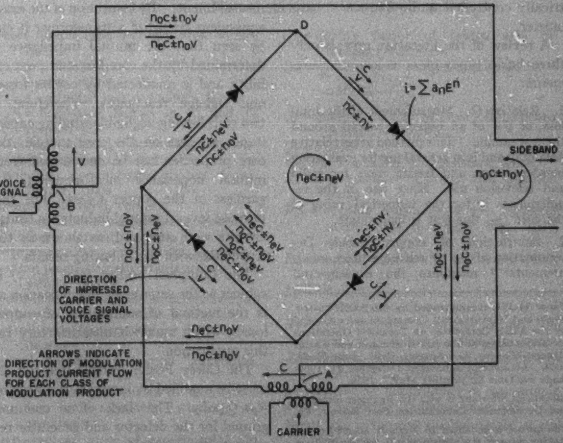


Fig. 13. Flow of modulation products within a ring modulator

Modulators ultimately reach a limit of usability because of the difficulties and cost of the associated filter circuitry. Further avenues are opened with the phase modulators, and it is hoped that this may be the path of further exploratory work.

References

1. COPPER-OXIDE MODULATORS IN CARRIER TELEPHONE SYSTEMS, R. S. Caruthers. *AIEE Transactions (Electrical Engineering)*, vol. 58, June 1939, pp. 253–60.

2. APPERT MODULATOR, Lenkurt Electric Company, Inc. United States Patent No. *2,545,250*, Mar. 13, 1951.

3. THE N-1 TELEPHONE CARRIER SYSTEM: OBJECTIVES AND TRANSMISSION CHARACTERISTICS, R. S. Caruthers. *Bell System Technical Journal*, New York, N. Y., Jan. 1951, pp. 1–32.

◆

No Discussion

Measurement of the Quality Factor of Inductor Cores

CHANDLER STEWART
NONMEMBER AIEE

Requirement for a New Test Method

THE Engineer Research and Development Laboratories (ERDL), Fort Belvoir, Va., have encountered, in connection with investigations into the effect of electrical component characteristics on the performance of various army equipment, a requirement for information on the electrical properties of low-loss core materials at low flux densities.

In particular, the effect of temperature on permeability and quality factor Q_m is required for application to adjustable differential transformers with close phase-shift tolerances. In addition to this information, the effects of temperature on saturation characteristics are required for application to a low phase-shift electrically controlled audio-frequency variometer.

A review of the literature reveals only three basic types of core loss measurements:

1. Relative Q. Measurement of the total loss, or Q_T, of an inductor wound around the test sample. This method gives relative data only and is in general use for evaluating powdered iron and ferrite cores for radio and television use. Since one of the requirements is for the numerical value of the core Q_m, this test is inadequate.

2. Subtraction of copper losses. Determination of Q_m by self-impedance measurements[1–4] magnifies the experimental errors, as shown in the Appendix. In cases of very high Q_m the results may be worthless. Also, this method imposes the practical difficulty of winding a very large amount of wire on each test sample.

3. Wattmeter. In the Epstein test[5] the core loss is indicated on a low power factor wattmeter. The current coil is in series with the primary winding of the test transformer and the potential coil is across the secondary winding. Such a method gives satisfactory results only within the calibrated frequency range of the wattmeter and then only for relatively low values of Q_m.

Determining Core Properties by Mutual Impedance Measurement

Because of these limitations it became necessary to devise a new method of measurement. By inspection of the exact equivalent circuit of a transformer it can be seen that the mutual impedance is determined by the core loss and permeability and is unaffected by copper losses and leakage reactances. Therefore, if two coils having negligible winding capacitance are put on the core sample, the core properties can be determined from mutual impedance measurements, regardless of the copper losses. For this purpose several mutual inductance bridge circuits were analyzed, as shown in the Appendix, with the following results.

The Campbell bridge[6] (see Fig. 1) is subject to the same error magnification as is the method of subtraction of copper losses, which makes it unsatisfactory for this application.

The Carey Foster bridge[4] (see Fig. 2) must be modified by insertion of a resistor $R_P > Q_m \omega L_P$. The lack of a common ground for the detector and generator requires either a bridge transformer, involving consideration of magnetic shielding and interwinding capacity, or a Wagner ground, with its attendant operating difficulties. A separate measurement of primary impedance is required, which limits the over-all accuracy, especially when the effect of direct current on the core is being observed. For these reasons this bridge was not used.

The Felici bridge[6] must be modified to provide a loss adjustment, such as shown in Fig. 3. Also, a bridge transformer or Wagner ground is required. The generator, bridge transformer, and core sample test transformer must all be well shielded to avoid stray coupling with the variable inductor, which must necessarily be unshielded and develop a large external field. Also, the variable inductor must be carefully located to avoid eddy-current losses and pickup of noise from power-line harmonics. Similarly constructed and with similar limitations is the bridge described by Campbell.[7]

The ERDL Mutual Inductance Bridge

To avoid these limitations a new bridge circuit, shown in Fig. 4, was developed. An analysis of this circuit is given in the Appendix. It was found convenient to employ ring samples about 1 to 2 inches in diameter. The biasing field is

$$H_{DC} = 0.495\, I_{DC}(N_1 - N_2)/l_1 \text{ oersteds}$$

The a-c field is

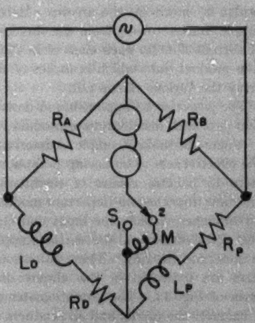


Fig. 1. Campbell bridge

Paper 54-291, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted December 21, 1953; made available for printing May 7, 1954.

CHANDLER STEWART is with the Engineer Research and Development Laboratories, Fort Belvoir, Va.

The author is indebted to Herbert Marsteller for making most of the measurements, and to William Kohmann for checking the mathematical derivations.

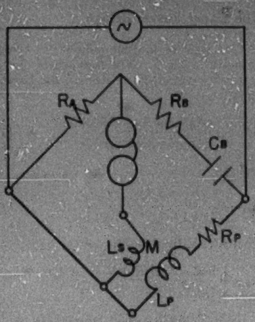


Fig. 2. Carey Foster bridge

$H_{AC}=0.495\,I_P(N_1-N_2)/l_1$ oersteds

The core properties are given by

$\mu=l_1/\omega^2(C_1+C_2)3.19\times10^{-8}\,N_1N_2a$

and

$$Q_m=\left(\frac{1}{C_1}+\frac{1}{C_2}\right)/\omega R_1\left(1-\frac{D_1}{\omega R_1C_2}\right)$$

This bridge offers the following advantages:

1. So long as the Q of C_1 is at least three times Q_m (easily achieved in high-quality mica or polystyrene capacitors), and C_1 is several times as large as C_2, there is negligible magnification of errors.

2. Accuracy is limited only by the calibrations of resistors and capacitors, which are generally superior to those of inductors, especially if the latter employ iron cores.

3. Since no bridge transformer or inductors are used, there is no stray magnetic induction problem or noise pickup problem. As in any bridge, however, tests at extremely low levels may dictate the use of a step-up input transformer in the null detector to bring the null signal above shot noise in the first stage of the null detector.

4. The windings on the core sample may be very skimpy and easily put on by hand. For a 1,000-cycle test 100 turns would be typical.

5. The core sample winding wire gauge may be selected purely for mechanical convenience, since the resistance of the wire is not involved.

Measurements Results

As a rough check on the over-all operation of the ERDL mutual inductance bridge, the results were compared with data from other sources as in the following.

Molybdenum permalloy dust core, $\mu=125$, Q at 1,000 cycles:

1. Computed from data by Legg and Given,[8] $Q_m=1{,}020$

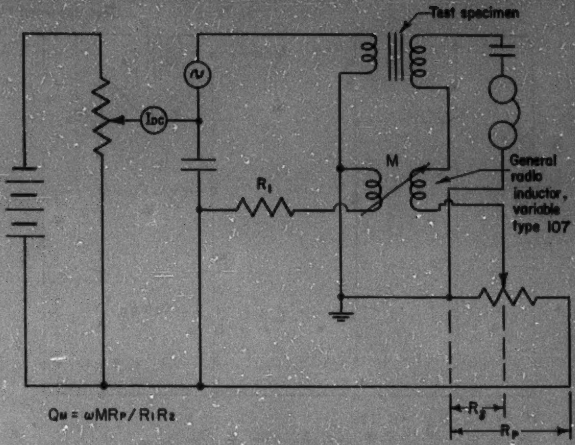


Fig. 3. Modified Felici bridge

2. Reported by Arnold Engineering Company, Marengo, Ill., $Q_m=600$ to 800

3. ERDL mutual inductance bridge, $Q_m=730$

General radio variable inductor, type *107*, Q at 1,000 cycles:

1. Theoretical Q, neglecting eddy currents in metal parts and neighboring metal, $Q_m=\infty$

2. ERDL mutual inductance bridge, $Q_m=800$

(It was noted that this Q depended upon the location of the inductor, and it is therefore assumed that the result would have been much higher if the measurement were made with the inductor located in a position further removed from surrounding metallic objects.)

Table I summarizes the results of measurements on a number of core specimens.

Conclusions

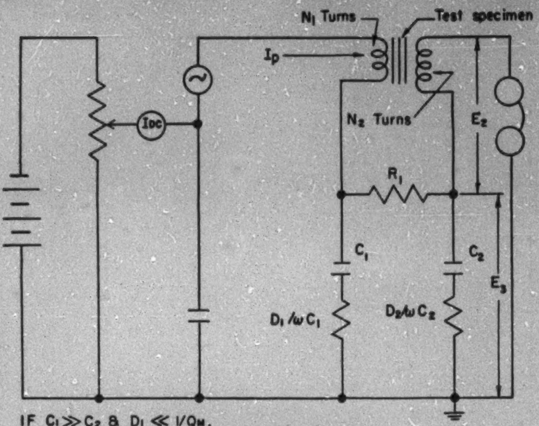
Measurement of magnetic core quality factor by the mutual impedance bridge described here offers more accuracy and

---

Nomenclature

a = core sample cross-sectional area, square inches
B_m = mutual susceptance of test transformer, mhos
C_1, C_2, C_B = bridge arm capacitors, farads
D_1, D_2 = dissipation factors of C_1 and C_2 respectively
E_2 = voltage induced in secondary winding
E_4 = voltage across C_2
G_m = mutual conductance of test transformer, mhos
H_{AC} = a-c field level, oersteds
H_{DC} = core sample biasing field, oersteds
I_{DC} = biasing current, amperes
I_P = primary alternating current, amperes
l_1 = mean magnetic path length around core sample, inches
L = self-inductance of test inductor, henrys
L_D = bridge arm inductance, henrys
L_P = test transformer primary inductance, henrys
M = test transformer mutual inductance, henrys
N_1, N_2 = number of turns on primary winding and secondary winding respectively
P_0 = percentage error of experimental observation
P_R = percentage error of computed result
Q_D = quality factor of bridge arm inductance $=\omega L_D/R_D$
Q_m = effective core sample quality factor $=\omega M/R_m=\left|\frac{B_m}{G_m}\right|$
Q_P = quality factor of test transformer primary $=\omega L_P/R_P$
$Q_T=\omega L/R_T$
R_A, R_B, R_1 = bridge arm resistors, ohms
R_D = a-c resistance of bridge arm inductor, ohms
R_{CU} = winding resistance of test inductor attributable to copper losses only, ohms
R_m = mutual resistance of test transformer, ohms
R_P = total effective resistance of transformer primary winding, ohms
R_T = total effective resistance of test inductor including coupled core losses
Y_m = mutual admittance of test transformer $=G_m+jB_m$, mhos
μ = effective a-c permeability of core sample
ω = test frequency, radians per second

Note: Q_m is the real quality factor of the material if the core has no air gap. μ is the real permeability of the material if the core has no air gap.



IF $C_1 \gg C_2$ & $D_1 \ll 1/Q_M$,
$Q_M = (\frac{1}{C_1}+\frac{1}{C_2})/\omega R_1 (1-D_1/\omega R_1 C_2)$ & $M = 1/\omega^2 (C_1+C_2)$

Fig. 4. ERDL mutual inductance bridge

convenience than other methods in general use.

Appendix

Core Testing by Self-Impedance Measurements

The core loss is found by subtracting the losses of the coil alone from the losses of the coil with the core in place.[1-4] The quality factor can be obtained from impedance bridge measurements by

$$\frac{1}{Q_m}=\frac{1}{Q_T}-\frac{R_{CU}}{\omega L} \tag{1}$$

The percentage error in such a measurement can be estimated from statistics[9,10] as follows

$$\frac{P_R}{Q_m}=\left[\frac{P_0^2}{Q_T^2}+\frac{(P_0R_{CU})^2}{(\omega L)^2}\right]^{1/2} \tag{2}$$

Since

$$Q_m>>Q_T \tag{3}$$

equation 2 can be approximated as

$$P_R\cong\sqrt{2}Q_mP_0/Q_T \tag{4}$$

Since Q_m is often, in practice, from 10 to 100 times Q_T, the self-impedance measurements method of Q_m determination seriously magnifies the experimental errors. Likewise, the method imposes the inconvenience of requiring a large winding to maintain a high value of Q_T.

The Campbell Bridge

For the Campbell mutual inductance bridge,[8] shown in Fig. 1, the following balance equations apply

$$M=\pm\frac{R_AL_P-R_BL_D}{R_A+R_B} \tag{5}$$

and

$$R_m=\pm\frac{R_AR_P-R_BR_D}{R_A+R_B} \tag{6}$$

Then, by definition of Q_m and from equations 5 and 6

$$Q_m=\frac{1-R_BL_D/R_AL_P}{1/Q_P-R_BL_D/R_AL_P\times 1/Q_D} \tag{7}$$

Since the magnetic loss must be less than the total loss

$$Q_m>Q_P \tag{8}$$

Assuming that experimental errors occur only in the values of L_D, L_P, Q_P, and Q_D, it is found that a statistical analysis[9,10] of the probable error of equation 7 gives the following

$$\frac{P_R}{P_0}=\frac{\sqrt{(Q_m-Q_P)^2+(Q_m-Q_D)^2}}{Q_D-Q_P} \tag{9}$$

Inspection of equation 9 shows that the resulting error will be prohibitive unless either Q_D or Q_P is reduced well below the value of the other. Reducing Q_P is the more convenient, since this also reduces the size of the primary winding and consequently reduces the work involved in preparing each test sample. If Q_P is reduced sufficiently, the following approximation of equation 9 can be based on equation 8

$$\frac{P_R}{P_0}\cong\sqrt{2}\frac{Q_m}{Q_D} \tag{10}$$

This equation is similar to equation 4 and hence the error magnification in the Campbell bridge is similar to that incurred in self-impedance tests.

The Carey Foster Bridge

This type of mutual inductance bridge[8] is shown in Fig. 2. For cases where mutual resistance cannot be neglected, the balance equations are

$$M=R_A\frac{(R_A+R_B)L_P+R_P/\omega^2C_B}{(R_A+R_B)^2+1/\omega^2C_B^2} \tag{11}$$

and

$$R_m=R_A\frac{(R_A+R_B)R_P-L_P/C_B}{(R_A+R_B)^2+1/\omega^2C_B^2} \tag{12}$$

from which, by definition of Q_m

$$Q_m=\frac{1+\omega C_B(R_A+R_B)Q_P}{\omega C_B(R_A+R_B)-Q_P} \tag{13}$$

If the errors in a determination of Q_m by this bridge are attributed to errors in measuring Q_P, the percentage error can be statistically analyzed[9,10] with the following result

$$\frac{P_R}{P_0}=Q_mQ_P\frac{(1+Q_m^{-2})}{(1+Q_P^2)} \tag{14}$$

Inspection of equation 14 shows that the errors are greatly magnified unless

$$Q_P<1/Q_m \tag{15}$$

which requires that R_P be enlarged artificially by an additional bridge arm. Under such conditions, the Carey Foster bridge should be satisfactory for magnetic core testing. Its accuracy is limited by the errors in measurement of L_P and by the stability of L_P. It was not used in this investigation, however, because of problems arising from inability to connect the generator and detector to a common ground.

Modified Felici Bridge

The Felici bridge as described in the literature[8] is not generally capable of an exact balance since there is no provision for loss compensation, there being only one bridge arm. A modified form of this bridge, incorporating such an additional adjustment, is shown in Fig. 3. For this circuit, when

$$R_P>>R_1 \tag{16}$$

the mutual inductance of the transformer equals that of the variable inductor, and

$$Q_m=\omega MR_P/R_1R_2 \tag{17}$$

Statistical analysis[9,10] of this equation (under the assumption that ω, R_P, R_1, and R_2 can be read much more accurately than M) shows no error magnification. The greatest practical source of error arises from the large external field set up by commercially available variable inductors, which induces eddy currents in neighboring metallic objects in an unpredictable manner. An additional disadvantage is the extremely strong noise signal picked up by the variable inductor, mainly from power-line harmonics.

ERDL Mutual Inductance Bridge

This circuit, shown in Fig. 4, was developed for this investigation.

Table I. Measurements of Magnetic Properties of Cores*

Material and Sample No.	Test Method† at +70 Degrees Fahrenheit	A-C Level H_{AC}, Oersteds $\times 10^3$	Dissipation Factor $(1/Q_m)$ $\times 10^3$ at 1,000 Cycles per Second $(H_{DC}=0)$	Normal Permeability $\mu (H_{DC}=0)$	D-C Biasing Field H_{DC} to Reduce μ 50 Per Cent, Oersteds	Dissipation Factor $(1/Q_m)$ $\times 10^3$ at 50-Per-Cent μ $(\mu_\Delta = \mu/2)$§	Description of Material
Ceramic *B*	*D*	0.3	4.6	80			toroid
Ceramic *E*	*D*	0.2	7.4	570	>1.0	>7.4	toroid
Ceramic *G*	*A*	0.3	9.3	400			toroid
Ceramic *H*	*D*	0.7	10.0	485	>1.0	>10.0	toroid
Ceramic *J*	*D*	0.03	6.6	110	>2.5	>17.0	toroid
Ceramic *F4S*	*D*	0.09	6.1	430	2.2	5.8	toroid
Ceramic *F-15S*	*D*	10	7.3	145	>2.0	>14.0	toroid
Ceramic *F-27S*	*D*	0.04	14.0	450			toroid
Ceramic *F-73* (Lavite)	*D*	0.3	12.8	370	0.4	5.0	toroid
Ceramic *5N*	*A*	1.0	6.3	310			*E*-type core with air gap
Powdered iron *M*	*A*	2.6	14.3	6.1	>8.0		pot core with air gap
Ferrite *3M*	*D*	0.15	2.0	1,200			toroid
Ferrite *3C*	*D*	0.2	1.4				toroid
Ferrite *3CA*	*D*	0.15	1.5	870			toroid
	D‡	0.15	1.0	940			
Ferrite *3CB*	*D*	0.15	1.6	1,350			toroid
	D‡	0.15	1.1	1,420			
Ferrite *3CF*	*D*	0.21	5.2	885			toroid
Ferrite *3CG*	*D*	0.21	5.3	700			toroid
Ferrite *3CM*	*D*	0.37	1.3	850			toroid
Supermalloy *A*	*B*	0.15	156.0	44,000	0.01	160	toroid-tape wound; tape thickness 1 mil
	B‡	0.15	67.0	14,000			
	A	0.5	154.0				
Supermalloy *B*	*B*	0.2	147.0	34,000			toroid-tape wound; tape thickness 1 mil
	B‡	0.2	91.0	19,000			
Supermalloy *C*	*B*	0.08	455	41,000			toroid-tape wound; tape thickness 2 mils
	B‡	0.08	185	15,000			
Mo-Permalloy *A*	*B*	0.44	44.5	11,300	0.06	26	toroid-tape wound; tape thickness 1 mil
	B‡	0.44	41.6	12,000			
Mo-Permalloy *B*	*B*	0.17	52.6	16,000			toroid-tape wound; tape thickness 2 mils
	B‡	0.17	52.6	16,000			
Mo-Permalloy *D*	*D*	0.4	28.3	5,000			toroid-tape wound; tape thickness 1 mil
Deltamax	*D*	0.09	7.4	495			toroid-tape wound; tape thickness 1 mil
Deltamax *A*	*D*	0.11	9.2	475			toroid-tape wound; tape thickness 1 mil
Deltamax *C*	*D*	0.10	7.4	415			toroid-tape wound; tape thickness 1 mil
Deltamax *C*	*D*	0.16	8.4	424			toroid-tape wound; tape thickness 1 mil
Deltamax *C*	*D*	2.0	17.3	480			toroid-tape wound; tape thickness 1 mil
	D‡	2.0	21.2	385			
Deltamax *D*	*D*	0.13	17.6	615			toroid-tape wound; tape thickness 1 mil
Deltamax *D*	*D*	2.0	36.7	700			toroid-tape wound; tape thickness 2 mils
	D‡	2.0	47.0	595			
Mo-Permalloy *TC-0*	*D*	1.3	1.5	104			toroid-powder core
Mo-Permalloy *TC-0*	*D*	1.6	1.7	113			toroid-powder core
	D‡	1.6	1.7	114			
Powdered Iron *PY 17A*	D	3.4	1.9	33			toroid

* All cores assumed to have no residual magnetism.

† Test method *A*, Maxwell bridge; *B*, Owen bridge; *D*, ERDL mutual inductance bridge.

‡ Measurement made at −60 degrees Fahrenheit.

§ μ_Δ = incremental permeability.

$$E_3=\frac{(D_1-j)(D_2-j)I_P}{\omega^2 C_1C_2[R_1+(D_1-j)/\omega C_1+(D_2-j)/\omega C_2]} \tag{18}$$

$$E_2=-I_P/Y_m \tag{19}$$

For a null balance

$$E_2=E_3 \tag{20}$$

Thus, equating equations 18 and 19

$$Y_m=G_m+jB_m$$

where

$$G_m=\frac{[R_1\,\omega^2C_1C_2(1-D_1D_2)-\omega C_1D_1(1+D_2^2)-\omega C_2D_2(1+D_1^2)]}{1+(D_1D_2)^2+D_1^2+D_2^2} \tag{21A}$$

$$B_m=-\frac{\omega(C_1+C_2)(1-D_1D_2)+R_1\omega^2C_1C_2(D_1+D_2)+\omega C_1D_2(D_1+D_2)+\omega C_2D_1(D_1+D_2)}{1+(D_1D_2)^2+D_1^2+D_2^2} \tag{21B}$$

Since

$$D_1<<1 \tag{22}$$

and

$$D_2<<1 \tag{23}$$

it is found that

$$Y_m\cong(R_1\omega^2C_1C_2-\omega C_1D_1-\omega C_2D_2)-j[\omega C_1+\omega C_2+R_1\omega^2C_1C_2(D_1+D_2)] \tag{24}$$

From equation 24, by definition

$$Q_m=\left|\frac{B_m}{G_m}\right| \tag{25}$$

$$Q_m=\frac{\omega C_1+\omega C_2+R_1\omega^2C_1C_2(D_1+D_2)}{R_1\omega^2C_1C_2-\omega C_1D_1-\omega C_2D_2} \tag{26}$$

If

$$C_1>>C_2 \tag{27}$$

equation 26 can be approximated from equations 22 and 23 as

$$Q_m\cong\frac{1/C_1+1/C_2}{\omega R_1(1-D_1/\omega R_1C_2)} \tag{28}$$

Since in practice the following equation should be made

$$D_1<<\omega R_1C_2 \tag{29}$$

equation 28 can be very roughly approximated as

$$Q_m\cong 1/\omega R_1C_2 \tag{30}$$

from which $M\cong 1/\omega B_m$ can be further approximated as

$$M\cong 1/\omega^2\left[C_1+C_2+\frac{C_1(D_1+D_2)}{Q_m}\right] \tag{31}$$

In practice

$$Q_m>>1 \tag{32}$$

and from equations 22, 23, and 27 equation 31 can be approximated with good accuracy as

$$M\cong 1/\omega^2(C_1+C_2) \tag{33}$$

In an analysis of probable error[9, 10] in the value of Q_m computed from equation 28, D_1 can be assumed to be the exclusive source of experimental error, since ω, R_1, C_1, and C_2 are able to be determined much more accurately than D_1. This gives

$$\frac{P_R}{Q_m}=\frac{P_0D_1}{1+\frac{C_2}{C_1}} \tag{34}$$

Substituting equation 26 into equation 34 gives

$$P_R=P_0/\left(\frac{\omega R_1C_2}{D_1}-1\right) \tag{35}$$

Inspection of equations 30 and 35 shows that no appreciable magnification of errors occurs so long as

$$D_1<<1/Q_m \tag{36}$$

as shown also by equation 34.

References

1. High Frequency Measurements (book), A. Hund. McGraw-Hill Book Company, Inc., New York, N. Y., second edition, 1951, p. 375.

2. Ferromagnetic Spinels for Radio Frequencies, R. L. Harvey, I. J. Hegyi, H. W. Leverenz. *RCA Review*, Princeton, N. J., vol. 11, Sept. 1950, p. 356.

3. Radio Frequency Measurements (book), L. Hartshorn. John Wiley and Sons, Inc., New York, N. Y., 1940, pp. 200 and 231.

4. Electrical Measurements (book), F. K. Harris. John Wiley and Sons, Inc., New York, N. Y., 1952, pp. 390–94.

5. Standard on Measurement of A-C Core Loss and Permeability of Magnetic Materials. *ASTM Standard No. A343–49*, American Society for Testing Materials, Philadelphia, Pa., 1952.

6. Electronic Measurements (book), F. E. Terman, J. M. Pettit. McGraw-Hill Book Company, Inc., New York, N. Y., 1952, p. 114.

7. On the Use of Mutual Inductometers, Albert Campbell. *Proceedings*, Physical Society, London, England, vol. 22, July 1910, pp. 207–19.

8. Compressed Powdered Molybdenum Permalloy for High-Quality Inductance Coils, V. E. Legg, F. J. Given. *AIEE Transactions*, vol. 59, 1940, pp. 865–72.

9. Chemical Computations and Errors (book), T. B. Crumpler, J. H. Yoe. John Wiley and Sons, Inc., New York, N. Y., 1940, pp. 177–78.

10. Electrical Engineers' Handbook, H. Pender, W. A. Del Mar. John Wiley and Sons, Inc., New York, N. Y., fourth edition, 1949, pp. 1–16.

No Discussion

Radio Transmission of Narrow-Band Mobile Radio Systems at 40 Megacycles

C. R. KRAUS
MEMBER AIEE

W. G. CHANEY
NONMEMBER AIEE

A. T. STEELMAN
ASSOCIATE MEMBER AIEE

MOBILE radio equipment manufactured in the period from 1946 to 1950 was designed for a 120-kc channel allocation in the 40-megacycle (mc) band. The receivers used with these systems had poor selectivity characteristics by today's standards and are now referred to as "extra-wide-band" receivers. As mobile radio service expanded and the demand for frequency allocations increased, the inefficiency of this bandwidth assignment became evident. Accordingly the channel allocation was reduced to 40 kc, thus tripling the available number of frequency assignments. Initially, the Federal Communications Commission made assignments on an alternate-channel basis so as to allow 80-kc frequency spread between stations in the same geographical area. However, by 1951 most manufacturers were supplying equipment capable of working in a 40-kc channel assignment without being affected by systems on adjacent channels. Currently, receiver design has progressed to the point where 20-kc channel allocations, split-channel operation, appear to be feasible, and units of this type have been manufactured and placed in service.

Technical literature, based largely on laboratory tests and theoretical computations, has compared favorably the performance of the narrow- and wide-band sets. Nevertheless, in practice, the introduction of narrow-band sets into existing systems comprised of extra-wide-band equipment caused immediate adverse comment from the operating personnel. In disagreement with the claims made by advocates of split-channel operation, surveys brought out the fact that the new units produced a reduction in operating radius and posed serious problems in the case of systems which cover large geographical areas and serve mobile units moving at high speeds on busy highways. It was obvious that the problem could not be side-stepped merely by avoiding the use of narrow-band sets. Instead, system design would have to compensate for the shortcomings of the new equipments which were forced into service by the need for additional frequency assignments. Unfortunately, neither the literature nor the early surveys gave any indication of the magnitude of the degradation introduced.

Because the field design engineer must deal both with new systems and with the expansion of existing systems, it was imperative that quantitative answers be found for two basic questions:

1. What is the comparative coverage of homogeneous extra-wide-band mobile radio systems and narrow-band mobile radio systems of the same power rating?
2. What is the comparative coverage of extra-wide-band and narrow-band units in a predominantly extra-wide-band system?

Because available technical literature did not answer these questions in terms

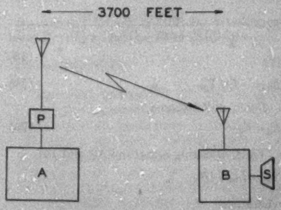


Fig. 1. Block diagram of test arrangements

A—Base transmitter arranged for control of modulation by means of deviation limiters
B—Mobile unit equipped with six different receivers and a coaxial switch to connect each receiver in turn to antenna and speaker
P—Variable pad to adjust radiated power of transmitter over a range of 80 db in 2-db steps
S—Speaker output of receivers

Paper 54-240, recommended by the AIEE Radio Communications Systems and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 30, 1954; made available for printing May 10, 1954.

C. R. Kraus, W. G. Chaney, and A. T. Steelman are all with the Bell Telephone Company of Pennsylvania, Philadelphia, Pa.

The authors wish to acknowledge the encouragement of A. Bailey of the American Telephone and Telegraph Company and V. A. Douglas of the Bell Telephone Laboratories in connection with the project. In addition, thanks are extended to D. O. Crofford, F. M. Hassenplug, and H. B. McCurdy as well as to other engineers of The Bell Telephone Company of Pennsylvania who participated in the test program. The courtesy and co-operation of Lt. Don Wagner, who so kindly made available facilities of the Pennsylvania State Police, is sincerely appreciated.

Table I. Receivers Used in Test Car

Receivers	Year Manufactured	Manufacturer
Extra-wide-band................	1946	A
Extra-wide-band................	1952 ..	A
Wide-band (40-kc system)......	1952 ..	A
Wide-band (40-kc system)......	1952 ..	B
Narrow-band (20-kc system)....	1952 ..	A
Narrow-band (20-kc system)....	1952 ..	B

of field conditions which include the presence of impulse noise from motor vehicle sources, it was decided that field tests were required. Tests were therefore conducted at Dover, Del., in April and May 1953 and at Gettysburg, Pa., in June 1953. The program of the tests was arranged primarily to provide answers to the two basic questions, but certain related factors such as the effect of frequency drift on the narrow-band receivers and the choice of frequency deviation for transmitters serving more than one type of receiver simultaneously were also investigated.

Subsequent to the work at Dover and Gettysburg, it was learned that a special subcommittee of the Joint Technical Advisory Committee had submitted a report on "Land-Mobile Channeling Arrangements." The report was primarily concerned with the 150-mc and 450-mc bands, making no mention of the 40-mc band. However, in spite of the frequency differences, the results are in close agreement on major points.

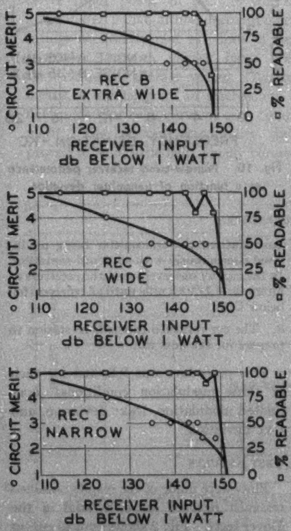


Fig. 2. Circuit merit and per-cent readability versus receiver input for three receivers

Description of Tests

TECHNICAL CONSIDERATIONS

The basic problem of the radio engineer designing a mobile radio system is to select equipment which will provide satisfactory coverage over a prescribed area. This requires that the received signal level at the mobile unit and at the base station be computed with reasonable accuracy. Two additional factors which may reduce the effectiveness of radio transmission must be taken into account: the amount of noise present at the receiver and the amount of distortion introduced by the system.

If purely random noise is encountered by a radio receiver, a relatively narrow bandwidth is acceptable and the receiver with the narrower bandwidth may give equal or better performance. In practice, a mobile receiver is exposed to impulse (nonrandom) noise arising in the vehicle or originating in near-by vehicles. Under such noise conditions with narrow-band receivers, if the bandwidth is reduced by the use of sharp cutoff filters the reception is impaired. Distortion is also introduced if a narrow-band receiver is used with a transmitted signal which exceeds its bandwidth capabilities.

The operating quality of a radio circuit, as determined by the received signal level and the noise-plus-distortion impairment, is generally defined in terms of circuit merit. The circuit merit rating system gives an arbitrary breakdown of five categories in which a particular transmission may be classified, defined as follows:

Circuit Merit 5 (CM5): No noise or distortion discernable; nearly perfect transmission.

Circuit Merit 4 (CM4): Some background noise; audio quality good.

Circuit Merit 3 (CM3): Obvious background noise and possible distortion; transmission understandable.

Circuit Merit 2 (CM2): Heavy noise; transmission barely usable.

Circuit Merit 1 (CM1); Noise controlling; transmission impossible.

Over ideal terrain, the quality of transmission is gradually reduced as the distance from the transmitter increases. Under such conditions, it would be difficult to define the boundaries of individual circuit merit limits. However, the shadowing effect of terrain features normally causes rather abrupt changes from one circuit merit rating to another and

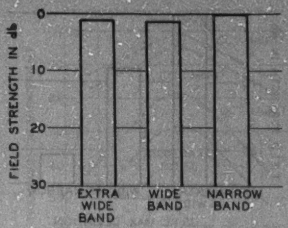


Fig. 3. Comparison of extra-wide, wide-, and narrow-band systems in the absence of impulse noise

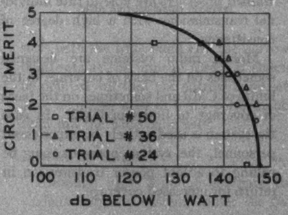


Fig. 4. Circuit merit versus receiver input for various trials in the absence of impulse noise

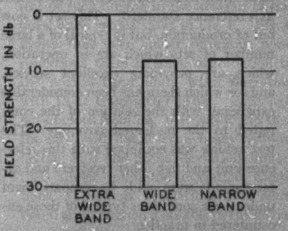


Fig. 5. Comparison of extra-wide, wide-, and narrow-band systems under moderate noise conditions (limit of CM3)

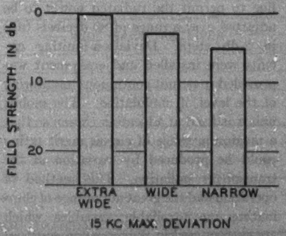


Fig. 6. Effect of mixing extra-wide, wide-, and narrow-band equipment in the absence of impulse noise

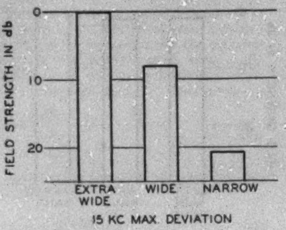


Fig. 7. Effect of mixing extra-wide, wide-, and narrow-band equipment under moderate noise conditions

trained observers are able to rate individual transmissions with a high degree of consistency.

Mobile radio systems are designed sometimes on the basis of coverage to the limits of CM3 and sometimes on the basis of coverage to limits of CM2. In case the system is used entirely by experienced personnel, the system operator may be willing to accept poorer transmission in return for increased coverage.

TESTING PROCEDURE

To evaluate receiver differences under controlled field conditions, the 1946 extra-wide-band Motorola receiver was chosen as a reference standard. This set is now out of production, but is typical of a large number of sets which are in operation, which are providing excellent service, and for which there has been considerable experience with the relation of the computed to the measured coverage. The test vehicle was equipped with this reference unit and five other receivers and arrangements were made to compare them under noise conditions typical of those encountered in practice.

TESTS AT DOVER

At Dover, a well-shielded base station transmitter was set up. Coaxial pads were inserted in the antenna transmission line to permit the radiated power to be adjusted over a range of 80 decibels (db) in 2-db steps. Deviation-limiting circuits were installed and equipment was provided to permit continuous monitoring of the level of modulation. The mobile unit was fixed at a location chosen so that a maximum range of circuit merit rating could be produced by variation of the transmitter radiation. This method of operation had the great advantage of eliminating uncontrollable variables which are encountered in the more conventional moving-car method of operation. Fig. 1 is a block diagram of the transmitter and receiver arrangements.

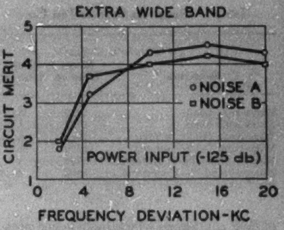


Fig. 8. Extra-wide-band receiver performance as a function of transmitter deviation

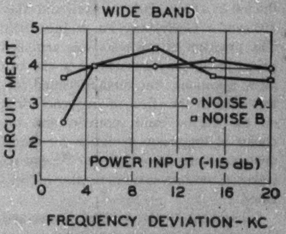


Fig. 9. Wide-band receiver performance as a function of transmitter deviation

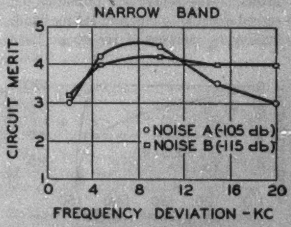


Fig. 10. Narrow-band receiver performance as a function of transmitter deviation

The test car was equipped with a coaxial switch to connect each of the six receivers in turn to the antenna. Loudspeakers were used for all monitoring in the test car; the proper speaker was selected automatically as each receiver was connected to the antenna. The car was equipped with rooftop, rear-mount, and cowl-mount antennas. The six receivers used are listed in Table I.

The quality of received signals was judged on a circuit merit basis by a skilled observer. Also, during most of the tests a second observer recorded the messages in order to derive a "per-cent readability" for each transmission. Tape recordings were made to aid in the analysis of the written data. Each trial of the test was run as follows:

1. At the transmitter, the carrier frequency and the deviation were adjusted. At the test car, the antenna, noise condition, and battery voltage were selected.
2. At the transmitter, the pad in the transmission line was adjusted so that the poorest receiver in the test car achieved a CM4 rating.
3. With the mobile antenna connected to the reference receiver, a message consisting of nine 2-digit numbers was transmitted. In the test car, one observer recorded his judgment of the circuit merit and a second observer copied the numbers of the message.
4. A similar transmission was made for each of the other receivers in turn.
5. The transmitted power was reduced 2 db by the pad and each receiver was again rated.
6. The process was repeated with power being reduced in 2-db steps until the best receiver received a CM1 rating.

Sixty-two trials were made in this manner to obtain sufficient data for different combinations of the variable factors.

TESTS AT GETTYSBURG

For the Gettysburg tests, the shielded transmitter was set up in the Pennsylvania State Police Barracks without pads in the antenna transmission line. U. S. Route 30, east of Gettysburg, is a straightaway route from this location over rolling terrain. The same test car which was used in the Dover tests was used in Gettysburg, but, in order to permit quick comparisons, only one each of the narrow-band, wide-band, and extra-wide-band receivers was used. Traffic conditions permitted a consistent road speed of 40 miles per hour. At this speed, it was possible to give circuit merit ratings to the three receivers every 1/10th mile. All runs were conducted as follows:

1. The transmitter deviation was set to the proper value.
2. The test car proceeded away from the base station, the transmitter being modulated continuously with standard sentences. In the car, an observer rated the receivers at intervals of 1/10th mile until all receivers fell below CM2.
3. The car returned to the base station to prepare for the next run.

This procedure was followed for each of three modulation conditions: non-limited modulation, 15-kc deviation, and 6-kc deviation.

TEST RESULTS

Plots were made of the results obtained on each receiver for each trial of the Dover tests. Fig. 2 shows typical curves for three of the receivers in only one of the trials. Here both circuit merit and per-

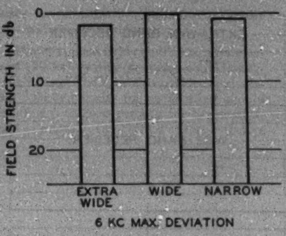


Fig. 11. Effect of mixing extra-wide, wide-, and narrow-band equipment in the absence of impulse noise

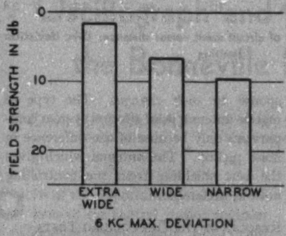


Fig. 12. Effect of mixing extra-wide, wide-, and narrow-band equipment under moderate noise conditions

cent readability are plotted against receiver input power in db. It will be noted in Fig. 2 that, although circuit merit ratings necessarily plot as points in step formation, smooth curves have been drawn through average values. By this means, a unique point of intersection is determined for each curve at each circuit merit. The per-cent readability data proved to be of little help, chiefly because of the relatively abrupt change from 100-per-cent to 0-per-cent readability.

Fig. 3, a composite of three different trials made during the study, compares the extra-wide-band system, the wide-band system, and the narrow-band system in the absence of impulse noise. The lowest field strength which gave CM3 performance in any system was assigned a zero db reference value. This figure bears out the manufacturers' claim that narrow-band systems compare favorably with the older systems under "quiet" conditions. Such conditions may be found at base station receiver locations, but are not typical of motor vehicle operation on modern highways.

The data showing circuit merit readings versus received power input taken for the reference receiver during the three tests previously described are given in Fig. 4. The variations between tests were small, indicating that the changes in the setup were minor during the period of time of the tests.

In the presence of impulse noise the comparisons in Fig. 3 are radically changed. On the basis of engineering experience, operation of a mobile unit in suburban areas on moderately busy highways was simulated by the introduction of sufficient impulse noise so that an increase of 12 to 15 db in the transmitter radiation with respect to quiet conditions was required in order to maintain CM3 transmission quality with the reference receiver. The results of three trials made with noise levels in this range are shown in Fig. 5.

Under the moderate noise conditions of Fig. 5, both the wide-band and the narrow-band sets are seen to be some 8 db less sensitive than the reference set. Although Fig. 5 shows nearly identical results with the wide-band and the narrow-band sets at the limit of CM3, it should be pointed out that the narrow-band sets gave a smaller area of CM4 performance than did the wide-band sets.

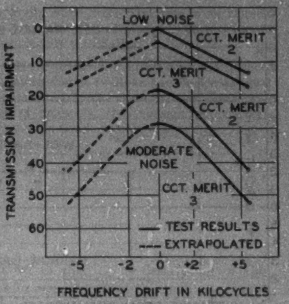


Fig. 13. Effect of off-frequency operation with narrow-band receiver

Mixing of Units

Although Figs. 3 and 5 compare the various complete systems under the two noise conditions, the results obtained when narrow-band units are added to an existing wide-band or extra-wide-band system are of especial concern. The effect of such mixing in the absence of impulse noise is illustrated in Fig. 6. Fig. 6 shows that even under optimum noise conditions there is some degradation in the narrow-band units when operating with 15-kc transmitter swing. Although not obvious from Fig. 6, it was found that the extra-wide-band sets actually perform better with the 15-kc deviation limiter than under unlimited deviation conditions. This may be attributed to additional amplification available in the devia-

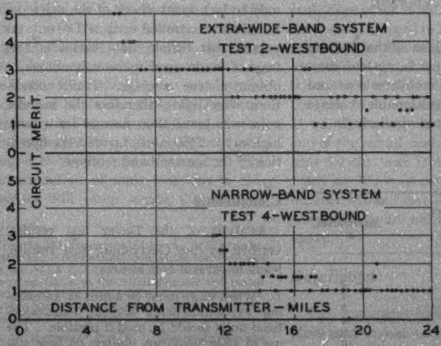


Fig. 14. Coverage tests of circuit merit versus distance: system comparisons

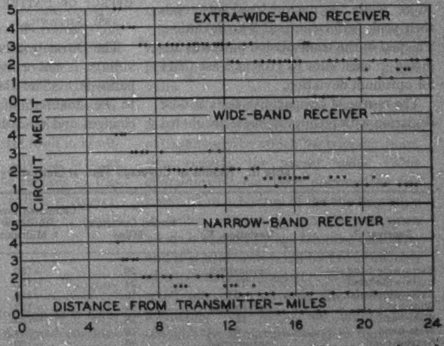


Fig. 15. Coverage tests of circuit merit versus distance: nonlimited deviation

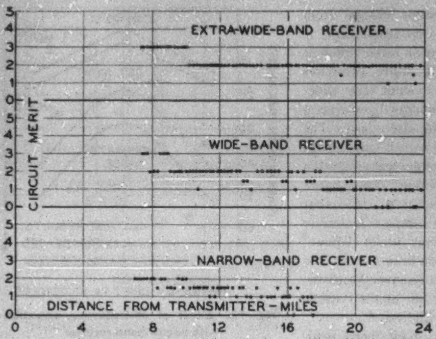


Fig. 16. Coverage tests of circuit merit versus distance: 15-kc deviation limiting

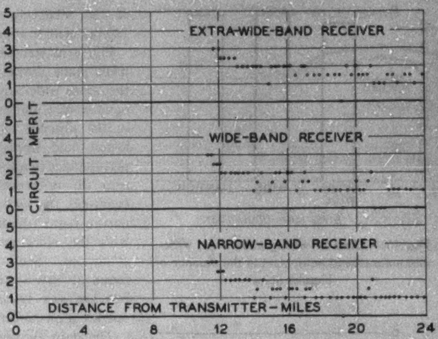


Fig. 17. Coverage tests of circuit merit versus distance: 6-kc deviation limiting

tion limiter which, combined with the limiting action, provides a higher average level of the modulating signal.

Under moderate noise conditions, the narrow-band unit operating in a wide-band or extra-wide-band system fares badly indeed, as may be seen in Fig. 7 which shows that a narrow-band set operating in a system employing 15-kc deviation suffers a degradation of 20 db as compared with the reference sets. With such a great penalty, mixing becomes entirely impractical.

When it is necessary to add narrow-band units to existing wide-band or extra-wide-band system, it is possible to reduce the penalty involved if a proper compromise frequency deviation is utilized. The effect of various transmitter frequency deviations on the extra-wide-band, wide-band, and narrow-band receivers is shown in Figs. 8, 9, and 10. The extra-wide-band and the wide-band sets perform best with a frequency deviation of 12 to 15 kc, while the narrow-band sets reach their peak performance at 8-kc deviation. The pad in the transmission path was adjusted in each case to give at least CM4 for optimum deviation.

The results of comparative tests at 6-kc deviation with no impulse noise are shown in Fig. 11. If Fig. 11 is compared with Fig. 6, it will be seen that in case a narrow-band fixed receiver at a quiet location is to be added to an existing system it might be advisable to reduce the deviation of associated mobile transmitters to approximately 6 kc.

The results of tests with 6-kc deviation under moderate noise conditions are shown in Fig. 12. A comparison of Figs. 12 and 7 shows that under moderate noise conditions the penalty for mixing equipments is definitely less if the deviation is reduced to 6 kc. However, it will be seen that the narrow-band set still suffers an 8-db degradation in comparison to the reference set.

Accurate frequency alignment is essential to the successful operation of narrow-band systems. The effect of operation slightly off frequency is shown in Fig. 13. A 2-kc departure (0.005 per cent) results in a 5-db penalty.

Effect of Battery Voltage and Antenna

The Dover test data show that battery voltage is not a factor in system comparisons. All of the mobile sets proved to be tolerant of reasonable voltage changes and all reacted similarly in response to such changes. The type of mobile antenna used affected system comparisons only because of the difference in noise pickup. The antenna which gives the best shielding from the controlling noise is preferable.

Results of the Gettysburg Tests

Samples of data taken in the Gettysburg tests are shown in Figs. 14 through 17. It will be seen that if radial coverage to a limit of CM2 is compared for the different combinations, the results are in fair agreement with the Dover conclusions. CM2 has been chosen for this comparison because the random fluctuations encountered in a moving vehicle make the determination of the exact limits of CM3 much less reliable. Table II lists the radial mileage, computed power in db, and the impairment relative to the extra-wide-band receiver with unrestricted modulation.

The Gettysburg tests were intended only to be a quick check of the more detailed and controlled tests at Dover. As pointed out earlier, data taken in this fashion would need to be voluminous to obtain a true average. Traffic conditions were light; therefore the impulse noise was lower than normal for a busy highway. This factor favored the operation of the narrow-band receiver.

Table II. Results of Coverage Tests at Gettysburg

Receiver	Deviation, Kc	Radial Coverage, Miles	Computed Power, Db Below 1 Milliwatt	Impairment, Db
Extra-wide-band	unrestricted	20	−130	0
Extra-wide-band	15	24	−132	−2
Extra-wide-band	6	19	−128	2
Wide-band	unrestricted	14	−122	8
Wide-band	15	18	−126	4
Wide-band	6	16	−124	6
Narrow-band	unrestricted	12	−120	10
Narrow-band	15	10	−116	14
Narrow-band	6	14	−122	8

Conclusions

Analysis of the Dover test results, verified by the Gettysburg test results, leads to several conclusions:

1. When units must be added to an existing extra-wide-band system, extra-wide-band units should be used if possible. The use of 15-kc deviation limiters on base station transmitters, which is now a Federal Communications Commission requirement,

causes no transmission impairment to the extra-wide-band equipment.

2. Under typical field conditions, a narrow-band (20 kc) system provides considerably less coverage than an extra-wide-band system. Design limits for application should include a 6 to 8-db penalty for the narrow-band system. In certain representative cases this will mean a reduction of 30 per cent in radial coverage.

3. If narrow-band units must be used in extra-wide-band systems it is advantageous to reduce the base station deviation. If the base station transmitter deviation is permitted to remain at 15 kc the penalty for the narrow-band units may be as high as 15 to 20 db, while if the deviation is restricted to 8 kc the degradation for narrow-band units is only 6 to 8 db.

4. If the regulatory authorities find it necessary to change to 20-kc bandwidth assignments in order to increase the available number of channels, the redesign of existing extra-wide-band systems will be required.

No Discussion

A New 12-Element Automatic Oscillograph and Application on the Bonneville Power System

C. M. HATHAWAY
MEMBER AIEE

W. L. DAVIS
ASSOCIATE MEMBER AIEE

J. R. CURTIN
ASSOCIATE MEMBER AIEE

ONE of the important operating problems of the central station engineer is to analyze system faults, to find out what the faults were and how well the protective equipment functioned to clear them. These faults may occur anywhere on the system and at any time during the day or night. They may last only a fraction of a second, and the most important part may last less than 1/2 cycle or less than 1/120 second.

As just one illustration of the importance of accurate fault records, consider the use of fault current values to locate short circuits and flashovers on long transmission lines. An insulator flashover will produce a short circuit to ground and the current which flows is determined by the voltage, the line impedance per mile of line, and the distance from the generator to the short circuit. A knowledge of the fault current will, therefore, provide the operating engineer with information from which he can calculate the location of the fault. Since the fault current may last for only 1/120 to 1/10 second, high-speed recording equipment is required to supply this information. The determination of fault locations in this manner will often locate the fault to within one or two spans of a long transmission line.

The quantities which should be recorded during a fault often include line and phase voltages, line currents, neutral current, phase-sequence components of current, and power. The records must be oscillographic so that wave shapes and phase relations and small time intervals can be determined.

This paper describes a new automatic oscillograph designed to produce an oscillographic record of 12 quantities of voltage, current, and power during unanticipated faults. These quantities are recorded on a chart traveling at a speed of 12 inches per second. The oscillograph will wait indefinitely with no parts in motion until a fault occurs, and when a fault does occur it will start and come to full recording speed in 2 milliseconds (0.002 second) after the initiating quantity has exceeded the preset tripping value. After it has been started the oscillograph will continue to run until the fault has cleared. It will then stop, record the time on the end of the record, and reset itself in readiness for the next fault. It can take as many as 100 consecutive records in this manner without any attention whatever, whether they are seconds, days, or weeks apart.

Description

This automatic oscillograph is a complete unit in one case designed for permanent switchboard mounting with back-of-panel connections. The front view is shown in Fig. 1. A description of the major components of this instrument and how they function together to produce the results described follows.

STARTING RELAYS

It is frequently desirable to arrange the automatic oscillograph to start from abnormal values of any one of several quantities such as overload, undervoltage, and neutral current. As many as three starting relays can be installed in this instrument if desired, to provide for automatic starting from any three of the following quantities:

1. Undervoltage, single phase
2. Undervoltage, 3 phase
3. Overcurrent, single phase
4. Overcurrent, 3 phase
5. Neutral current
6. Positive- or negative-sequence current
7. Positive- or negative-sequence voltage

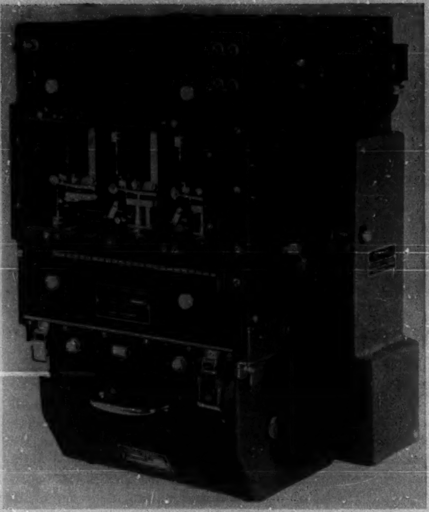
A starting relay is a device which measures the quantity by which it is controlled and trips whenever this quantity exceeds a definite preset value. For use in controlling a high-speed automatic oscillograph with an over-all starting speed of 2 milliseconds, it is evident that the tripping time of the starting relay must be less than 2 milliseconds by the time necessary to get the oscillograph in full-speed operation after the relay has tripped. The tripping time of the starting relay is about 1/2 millisecond after the initiating quantity exceeds the preset value.

For this application, the tripping time should be very nearly constant, regardless of the amount by which the tripping quantity exceeds the preset value. This characteristic has been obtained as follows: The armature of the starting relay is pulled upward by an electromagnet which is excited by the tripping quantity, and it is held downward by the attraction of a permanent magnet. When the upward attraction exceeds the downward attraction by even an infinitesimal amount the armature starts to move upward, and the first increment of motion breaks the contact between the armature

---

Paper 54-293, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 22, 1954; made available for printing May 10, 1954.

C. M. HATHAWAY and W. L. DAVIS are with the Hathaway Instrument Company, Denver, Colo., and J. R. CURTIN is with the Bonneville Power Administration, Portland, Oreg.



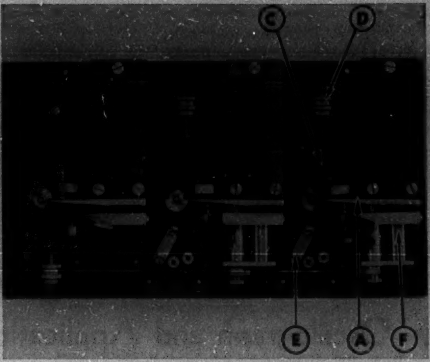


Fig. 1 (left). Twelve-channel automatic oscillograph

Fig. 2 (above). Starting relay panel

and the permanent magnet. When this contact is broken the attraction of the permanent magnet is greatly reduced and a large force difference is produced to move the armature. This arrangement provides a tripping time of about 1/2 millisecond when the initiating quantity exceeds the preset value by over 1 per cent.

Three relays are shown on the starting relay panel in Fig. 2, one for undervoltage, one for line current, and one for neutral current. In the line current and neutral current relays, the electromagnet which is energized by the initiating quantity is shown at the top. Directly below the electromagnet is the armature *A* which is held downward by the permanent magnet assembly *F*. The spring *C* aids the electromagnet in pulling the armature upward and increases the tripping speed. The tripping value can be adjusted accurately by means of the nuts *D* which control the spring tension. The starting coils are tapped at several points and the taps are brought out to the link switch *E* to provide the operator with a convenient means of changing the tripping value by known amounts. With the neutral current link in the position shown, this relay will trip at 0.25 ampere, and by moving the link the tripping value can be changed to 0.5 or 1.0 ampere.

The starting relay panel is mounted under the glass cover on the front of the oscillograph where it is shown in Fig. 1. A relay panel can be removed quickly after removing the cover and then removing four screws which hold it in place. All electric connections between the relay panel and the oscillograph are made by a multiple-link connecting strip so that it is convenient to remove the relay panel for changing the type of starting relays used.

RECORD MAGAZINE

The chart is contained entirely in the record magazine suspended below the oscillograph in Fig. 1. The chart is photosensitive paper, 10 inches in width, which is loaded into the record magazine in 200-foot rolls. As records are made the chart is transferred to a receiving spool which is also located in the record magazine. The record magazine can be removed whenever it is desired to remove the records or to reload it.

The record magazine is shown removed from the oscillograph in Fig. 3 and its cover is lifted off to show the supply spool supporting pin at *D* and the take-up spool at *E*. The chart is driven from one spool to the other by means of the rubber-faced drive roller *A*. When the cover is in place the chart is held against the drive roller by the pressure roller *B*, so that the chart is positively driven by turning the drive roller shaft.

CHART-DRIVE MECHANISM

When a fault is recorded the chart in the record magazine is driven from the supply to the take-up spool by turning the drive roller shaft. The shaft is driven by a governor-controlled motor which holds the chart speed at a constant value after the motor has come up to speed. When the record magazine is in place on the oscillograph, the gear *K*, Fig. 3, engages a gear in the oscillograph which is driven by the motor.

When the oscillograph is initiated by the tripping of a starting relay, the motor relay is closed and the motor is energized. The motor accelerates and comes to full operating speed in about 1/8 second.

The time required for the motor to reach full speed is much too long to permit recording of the first part of the fault, and in many cases the fault will not last as long as 1/8 second. To provide for acceleration of the chart to full speed in less than 2 milliseconds an auxiliary device called the initiating mechanism is coupled to the other end of the drive roller shaft by means of the coupling *L*, Fig. 3. The initiating mechanism releases the energy stored in a torsion spring to provide the required acceleration of the chart. So that it will be unnecessary to accelerate the entire 200-foot chart supply spool in $1^1/_2$ millisecond, the chart is brought from the supply spool around the idler roller *C*, Fig. 3. The idler roller is mounted on a pivoted carriage so that when the initiating mechanism demands a supply of chart the idler roller is pulled toward the drive roller. This shortening of the loop supplies chart while the supply spool is coming up to speed. After the supply spool comes up to speed and is able to supply chart as fast as it is needed, a spring returns the idler roller to its original position.

Both the chart-drive motor and the initiating mechanism drive the chart-drive roller through an overrunning clutch. During the first part of the record, while the motor is starting, the initiating mechanism drives the roller from the right, and the overrunning clutch on the left slips so that the initiating mechanism does not have to drive the motor. During the remainder of the record the overrunning clutch on the right slips and the motor does not have to drive the initiating mechanism. With this simple arrangement the motor will take over the driving of the chart just as soon as its speed slightly exceeds the speed of the initiating mechanism. The point on the record at which the drive is transferred from the initiating mechanism to the drive motor is undistinguishable, and the time axis is very nearly uniform after $1^1/_2$ milliseconds.

Initiating Mechanism

The initiating mechanism is shown in its position on the side of the case in Fig. 1. It is attached to the case by means of two screws and a cable jumper so that it can be quickly attached or removed. In applications which do not require 2-millisecond starting it is not required, but it can be attached in a few minutes when high-speed starting is needed. The initiating device removed from the oscillograph is shown in Fig. 4, in which its cover has been removed to show the internal mechanism. In Fig. 4, *D* is the starting spring and *U* a small motor which rewinds the starting spring after it has been used.

When the starting spring is rewound and ready to operate the spring is prevented from unwinding by the engagement of the pawl *E* with the sear *F*, and the sear is held against the pawl by the electromagnet *L*. Whenever the electromagnet is de-energized the pawl is released and the spring is allowed to unwind, and as it does so it drives the chart.

As the starting spring unwinds it pulls downward on the steel wire which is wrapped around a pulley attached to the starting spring and pawl. This pulls upward on a piston within the oil dash pot *K*. After the spring is released, it accelerates rapidly until the force exerted by the movement of the piston through the oil is equal to the force exerted by the spring on the wire around pulley *H*. The spring then continues to unwind at approximately uniform velocity. The acceleration of the spring to the condition of uniform velocity takes place in about 1 millisecond. Again in Fig. 4, *T* is a counter which is automatically advanced by means of a rotary solenoid at the end of each record so that an observer can tell how many faults have been recorded since the last inspection.

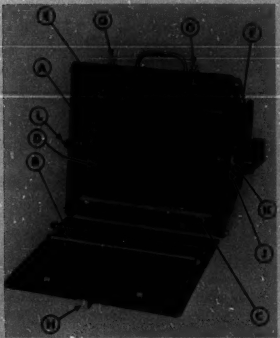


Fig. 3. Record magazine, cover removed

Timing Mechanism

The timing mechanism controls the length of the record after each initiation, and it can be set for records of 2 to 20 seconds duration. When it is set at 2 seconds, for example, the oscillograph will run for 2 seconds after an initiation and it will then stop and reset itself provided that the fault has cleared. If the fault has not cleared at the end of 2 seconds, the oscillograph will continue recording until it has cleared.

The spring-loaded starting relays will not reset themselves after the initiating quantity has returned to normal, hence it is necessary to perform this operation before the oscillograph is stopped and reset. This is done by the timing cam when it closes a pair of contacts which energize a resetting rotary solenoid on the back of the starting relay panel. After resetting the starting relay the cam continues rotation until it closes another set of contacts which reset the oscillograph. If the fault has not been cleared the starting relays will not remain reset and the oscillograph will not reset or stop recording.

Shutter Mechanism

The recording spots of light enter the record magazine through a shutter which is normally closed. When it is desired to take a record the shutter is opened in less than 2 milliseconds. At the end of the record the shutter is closed again by means of a rotary solenoid.

The shutter is opened by means of a spring which is released by an electromagnet in exactly the same manner as the starting spring is released to initiate the record. This electromagnet is normally energized, and is de-energized when a starting relay is tripped. Operation by de-energization is used rather than operation by energization because of difficulty in building up the electromagnet current through its inductance in a sufficiently short time.

Control Circuits

A complete description of the control circuits and of the sequence of operations involved in taking a record and in resetting the oscillograph would be long and tedious and is not necessary for a general understanding of the principles of operation. However, a brief outline of the control sequence will not be out of order.

The tripping electromagnets for the starting spring and the shutter are designed to be normally energized. These are energized through normally closed contacts on the starting relays. Tripping of any starting relay de-energizes these electromagnets, releasing the starting spring and the shutter. Operation of the shutter closes normally open contacts and energizes a relay which controls the chart-drive motor. This relay closes and locks closed through a holding circuit.

The starting spring begins to unwind immediately and continues for about 1/5 second. At the end of 1/5 second it closes and causes a pair of contact fingers in the initiating mechanism to initiate the rewinding cycle. The rewind motor then

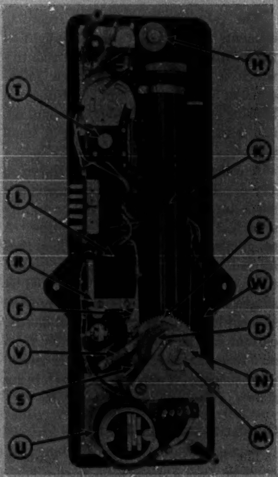


Fig. 4. Initiating device, cover removed

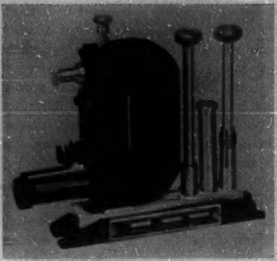


Fig. 5 (left). Bifilar galvanometer

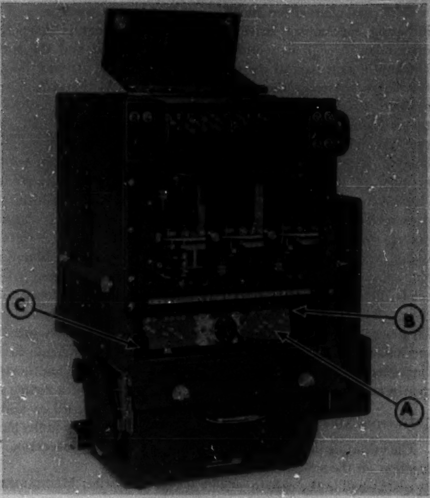


Fig. 6 (right). Automatic oscillograph with galvanometer and optical stage covers open

rewinds the starting spring and it is latched back in place.

The timing mechanism periodically tests the circuit to see if the fault has cleared by attempting to reset the starting relays. In about 1/2 second after it has succeeded in resetting the starting relays it closes a second pair of contact fingers which reset the shutter and stop the chart-drive motor.

After the chart-drive motor has been stopped and the shutter is closed a small lamp is flashed on which illuminates the face of a clock. By means of a small lens the clock face is projected onto the chart and is photographed there. The clock includes hour, minute, and second hands, an a.m.-p.m. dial, and a decade counter for counting up to 99 days. Blank space is provided on a card attached to the clock face for including the month and year, and station identification.

GALVANOMETERS

The galvanometers are the recording elements which move the spots of light on the record in accordance with the quantity which is being recorded. Each galvanometer is a complete unit consisting of a housing or cell box, a universal mounting, and a moving element carrying a small mirror.

A complete galvanometer is shown in Fig. 5. The moving element consists of a single turn of fine metallic ribbon tightly stretched between a pair of pole tips. A small mirror is cemented to the loop thus formed, and when current is passed through the loop the mirror is rotated by simple motor action. Since the moving element has a very low moment of inertia, current can be recorded accurately from zero to several thousand cycles per second in frequency. Twelve galvanometers are shown in the oscillograph in Fig. 6.

OPTICAL SYSTEM

The light for the recording spots is obtained from a single incandescent lamp (Mazda *1133* automobile headlight) which burns continuously. A part of the light is projected on each galvanometer by the optical stage shown at *A*, Fig. 6, which is an assembly of lenses, mirrors, and apertures. The entire optical stage assembly can be removed easily from the oscillograph by opening the hinged cover below the starting relay.

An aperture is interposed between the light source and each galvanometer, and the width of each recording trace can be individually adjusted by its aperture. These adjustments are accessible on the optical stage by opening the hinged optical cover, as shown at *B*, Fig. 6.

VIEWING SCREEN

A calibrated ground-glass viewing screen is located just above the optical stage cover, on which the positions and deflections of the recording spots can be observed at any time. The viewing screen is of great convenience in checking the condition of the recording spots and in making calibrations.

ACCESSORIES

The oscillograph galvanometers are used for measurement of voltage or current when connected to power system cur-

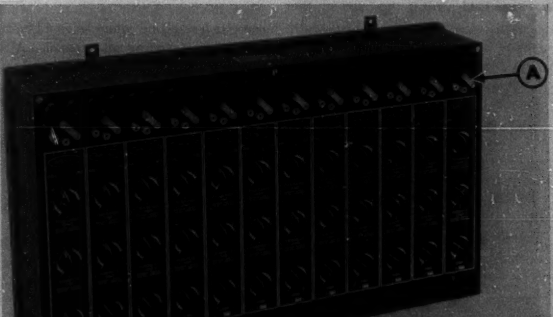


Fig. 7. Shunt resistor unit

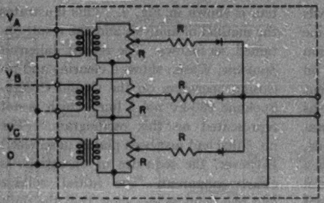


Fig. 8 (above). Circuit used for recording 3-phase voltages with one galvanometer

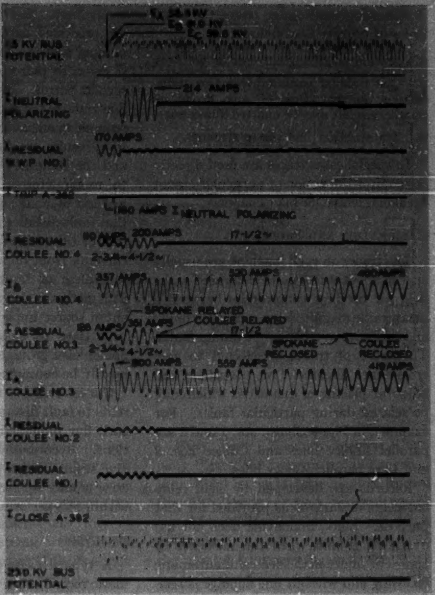


Fig. 9 (right). Automatic oscillogram of a clearing and reclosing operation on Grand Coulee-Spokane 230-kv line

rent transformers or potential transformers through a shunt resistor unit of the type shown in Fig. 7. The link *A* can be set for measurement of voltage or for measurement of current of 1 or 5 amperes nominal value.

Networks and filters for the relays are contained in separately mounted units. An accessory for placing time co-ordinate lines on the record can be added at any time. Galvanometers for measurement of single-phase or 3-phase power can be installed.

POWER SUPPLY

The 12-volt d-c model draws power for operation from a 100- to 140-ampere-hour battery which is kept in a continual state of charge by a power supply unit which provides current for stand-by operation of the lamp and holding coils and maintains a small charging current into the battery. An automatic voltage regulator and an output timer, which limits the maximum continuous recording period to a previously set value, are available as accessories to the power supply unit.

The 125- and 48-volt d-c models are designed to draw power from station service batteries. In these models the recording lamp receives its power from a separately mounted unit which contains a regulating-type transformer for providing essentially constant voltage to the lamp from 60-cycle-per-second 100- to 130-volt a-c. Sensing relays in the unit operate to supply lamp current through dropping resistors from the d-c line when the input alternating voltage fails or drops below the regulation range of the lamp transformer.

Automatic Oscillograph Practices on the Bonneville Power System

The Bonneville Power System is largely a network of long, heavily loaded high-voltage transmission lines designed to gain as much economy as possible by exploiting high-speed fault clearing, and automatic reclosing to increase stable power system limits.

One of the important functions of oscillographs installed on the system is the monitoring of over-all relay and circuit-breaker performance. The detection of relaying abnormalities and circuit-breaker troubles permits remedial measures before more serious failures occur.

Another important function is fault location. Prompt location and repair of fault damage, particularly on nonreclosable line faults, is vital to system reliability. Automatic oscillograph records afford the most practical means currently available for spotting maintenance crews close to trouble spots. Close location also contributes to the effectiveness, and reduces man-hours, of patrol crews.

Oscillograms of system swings are occasionally compared to network analyzer transient studies to verify the correlation between the two.

The oscillographs are sometimes temporarily installed in special locations for recording during staged system testing.

INSTALLATION PRACTICE

The newer Bonneville transmission switching stations are designed with panel space and current and potential terminal blocks to provide for oscillographs having about two traces per bus section and one or more traces per transmission line. These are usually centrally located in a terminal cabinet room beneath the control room relaying and control switchboards.

One set of phase and residual line currents per line terminal, one set of phase-to-neutral and residual potentials per transmission voltage bus section, and station polarizing currents are routed to the terminal blocks. Flexible connectors are provided to bring selected currents and voltages to the oscillograph-calibrating unit located inside the cubicle. Permanent connections are made between the calibrating unit and the oscillograph which is mounted on the outside of the cubicle.

TRACE ALLOCATION

The number of traces available to the application engineer is usually somewhat lower than might be desired, even in a 12-element oscillograph. To conserve elements, three phase-to-neutral voltages per bus are combined on one trace by means of half-wave rectifiers. See Fig. 8, which shows a schematic wiring diagram of this network. Phase identification is made possible by depressing one phase slightly. Calibration curves which show the effect of voltage change in each phase alone, and on all three phases together, are kept in file for reference when oscillograms are analyzed.

Following the voltage traces the usual order of priority is as follows:

1. Station polarizing currents or voltages.
2. One residual current per major transmission circuit. One of any group of parallel circuits may be omitted if necessary.
3. Transmission line phase currents.

In special cases traces are used directly to monitor the action of particular circuit breakers, relays, or pilot channels.

Calibration and Processing

The traces on each machine are usually spaced about equally and calibrated in the field for the best compromise between a measurable magnitude on minimum faults and excessive overlap on maximum faults. The order of traces is selected to give cross-chart alignment on potentials and adjacent placing of quantities which may be related during particular faults. For example, in Fig. 9 Coulee nos. 1 and 2 are parallel 115-kv lines and Coulee nos. 3 and 4 are parallel 230-kv lines.

Records are developed by field relay engineers in darkrooms provided for that purpose. Since Bonneville does not use overhead ground wires the incidence of transient faults with successful automatic reclosing and without line damage is relatively high. In these cases the records are processed and analyzed in a routine manner and the computed fault location is given to patrol crews for checking on regular patrol. In the case of permanent line outage the records are processed as soon as an engineer can get to the substation, results are compared by telephone, and the computed fault location is given to repair crews. The most successful method of analysis is based on the fact that fault distance is almost a linear function of the ratio of simultaneous residual currents from opposite ends of the line, regardless of fault type or impedance. This method is also substantially independent of source impedance except for faults very close to busses. Faults which do not involve ground are very rare and can readily be computed by other means.

Curves relating the ratio of fault currents to fault distances are maintained for each line under normal operating conditions. Records indicate that the curves give results which have had a maximum error of only 7 per cent and have averaged within $2\frac{1}{2}$ per cent of line length.

Analysis

A typical example of an actual automatic record taken on the Bonneville system is shown in Fig. 9. This illustrates the method of analysis and the result obtained, as follows: The record, taken at Spokane, Wash., shows a clearing and reclosing operation on the Grand Coulee-Spokane 230-kv no. 5 line, which is not represented on this oscillogram. (The no. 5 line had just been placed in service.) However, the sequence of operations is clearly apparent. The initial voltage and polarizing traces show that the fault was on *A*-phase near Spokane. The decrease in polarizing current and increase in parallel-line residual current show that the Spokane end opened in $2\frac{3}{4}$ cycles. The final drop in residual traces shows that Coulee cleared in $4\frac{1}{2}$ more cycles, and the pips on the polarizing and residual traces show that Coulee and Spokane both reclosed $17\frac{1}{2}$ cycles later. The first residual current pip lines up with spikes on both voltage traces.

Reference

1. An Automatic Oscillograph, C. M. Hathaway, R. C. Buell. *AIEE Transactions*, vol. 51, Mar. 1932, pp. 222–25.

◆

No Discussion

Type-N1 Carrier on Radio and Coaxial Cable

W. S. AMES
ASSOCIATE MEMBER AIEE

W. H. WEDEL
ASSOCIATE MEMBER AIEE

Synopsis: A practical illustration of the versatility obtainable in applying modern types of carrier telephone systems to radio facilities is given in an application of 48 channels of the Western Electric Company type-*N1* carrier to a 2-section radio system in an area served by The Pacific Telephone and Telegraph Company. The radio system, with the unusual feature of combining two types of radio equipments, has a Western Electric Company type-*TD-2* terminal at Los Angeles and type-*TE-1* terminal near Ventura, with one intermediate *TD-2* repeater at Oat Mountain, near Chatsworth, California.

A unique and economical aspect of the arrangement lies in the extension of the carrier systems from one of the radio terminals to the distant carrier terminal station through a 3-mile section of paired exchange cable and a 30-mile section of coaxial cable using intermediate *N1* carrier repeaters without departing from carrier frequencies. Another unusual feature of the arrangement consists of the method of positioning the carrier frequencies for radio transmission. Utilizing the availability of 12-channel *N1* carrier systems in two frequency allocations, low-group (44 to 140 kc) and high-group (164 to 260 kc), it was possible to combine two systems by means of filters to occupy a frequency range from 44 to 260 kc without additional modulators. Such a pair of *N1* systems with a total of 24 channels was then translated in a single step of modulation to a frequency range from 280 to 496 kc, and it was found that in accomplishing this single translation an *N1* carrier repeater, with modifications, performed admirably. Transmission performance of the over-all system is within the limits set for an all-cable *N1* carrier system of the usual type.

General

In 1950, a need suddenly arose for additional long-haul telephone circuits between Los Angeles and the coastal cities of Ventura and Santa Barbara. Means at hand which gave promise of providing service most quickly and economically to meet the toll circuit shortage were:

- Microwave repeater stations at Los Angeles and Oat Mountain (intermediate between Los Angeles and Ventura, dividing the distance into 28-mile and 38-mile paths).
- A location at Mills Road, near Ventura, suitable for a microwave terminal.
- Exchange cable pairs from Mills Road to the Ventura toll central office (3 miles).
- Coaxial cable between Ventura toll central office and Santa Barbara (29 miles).

Equipment obtainable to complete the facilities for the route consisted of 4,000-megacycle (mc) type-*TD-2* radio equipment for Los Angeles and Oat Mountain, 4,000-mc type-*TE-1* for Mills Road, 12-channel *N1* carrier terminals

---

Paper 54-214, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted February 23, 1954; made available for printing April 12, 1954.

W. S. Ames and W. H. Wedel are with the Pacific Telephone and Telegraph Company, Los Angeles, Calif.

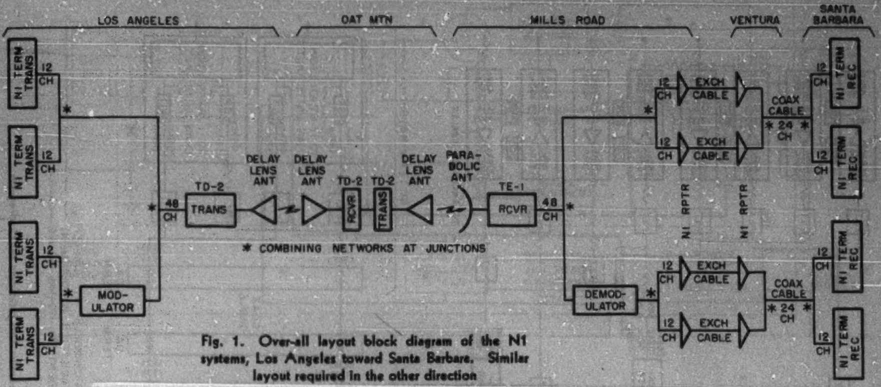


Fig. 1. Over-all layout block diagram of the N1 systems, Los Angeles toward Santa Barbara. Similar layout required in the other direction

for Los Angeles and Santa Barbara, and *N1* repeaters for use in the exchange cable and coaxial sections 1 through 6.

While the references give detailed descriptions of the *TD-2* and *TE-1* radio systems and the *N1* carrier telephone equipment, a brief mention of some of the more important features of these systems follows.

The *TD-2* microwave radio is designed for long-haul radio relay use to provide television network facilities or telephone message circuits in large quantities. The transmitted frequency is near 4,000 mc with a power output near 1/2 watt from an amplifier using *416A* triodes. Since the transmitter and receiver are designed for interconnection at 70-mc intermediate frequency (i-f) for use as a repeater, separate frequency modulation (FM) units are required at a terminal to associate the video band with the 70-mc i-f.

Designed for short-haul television relay use, the *TE-1* radio transmitter has a klystron output tube delivering somewhat less than 1/2 watt near 4,000 mc. The input of the transmitter and the output of the receiver are video frequencies.

Type-*N1* carrier is used extensively on nonloaded cable pairs to provide 12 voice channels per system. Both side bands are used for each channel, and the transmitted carrier frequencies serve as pilots to enable individual channel regulation in addition to group regulation. Two-line frequency allocations are available: 44 to 140 kc (low group) and 164 to 260 kc (high group). A complete *N1* carrier terminal uses transmitted frequencies in one group location with received frequencies in the other. Signalling is accomplished by a 3,700-cycle tone just above each 3-kc voice channel.

A block diagram of the layout (in one direction of transmission) is shown in Fig. 1, illustrating the application of four systems (48 channels) of *N1* carrier to the radio system, and showing the way in which the carrier systems are extended from the radio terminal at Mills Road via exchange cable and coaxial cable to Santa Barbara. As indicated in the sketch, the *N1* terminal transmitters at Los Angeles are combined in pairs with one pair being modulated upward to enable a combination of the four terminals before application to the radio transmitter. The four *N1* systems are separated at Mills Road to permit transmission through the exchange cable section, but they are again combined in pairs for transmission through the coaxial cable. Transmission in the opposite direction follows a similar arrangement.

Radio Sections

A general description of the radio portion of the system follows.

Los Angeles: Western Electric Company type-*TD-2* transmitter and receiver near 4,000 mc, with 10 by 10 by 10-inch delay lens horns, FM transmission with a power output of +27 decibels above 1 millowatt (dbm). Deviation was set for ±3 mc with all 48 channels of *N1* carrier applied.

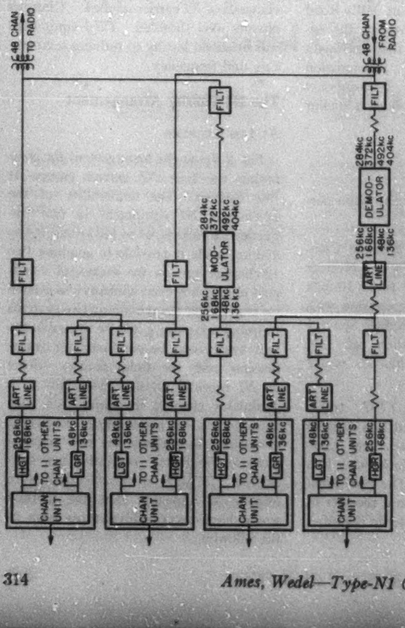
Oat Mountain: Type-*TD-2* transmitter and receiver toward Los Angeles and also toward Mill Road, near 4,000 mc, with 10 by 10 by 10-inch delay lens horns in each direction. The FM signal at 70 mc i-f is passed from a receiver in one direction to a transmitter in the other direction.

Mills Road: Western Electric Company type-*TE-1* transmitter and receiver near 4,000 mc, FM transmission toward the *TD-2* receiver and transmitter at Oat Mountain. *TE-1* power output is approximately +25 dbm. Because of the use of klystrons against the wave guide in the *TE-1*, distance between the equipment and the antennas is limited. At Mills Road 57-inch parabolic reflectors are aimed vertically through an alsynite window in the roof toward a 10 by 14-inch reflector elevated 60 feet above the earth. Deviation was set for ±3 mc with all 48 channels of *N1* carrier applied. Clamper circuits were disabled. *TE-1* equipment was modified locally to stabilize temperature and frequency.

The N1 Carrier Arrangement

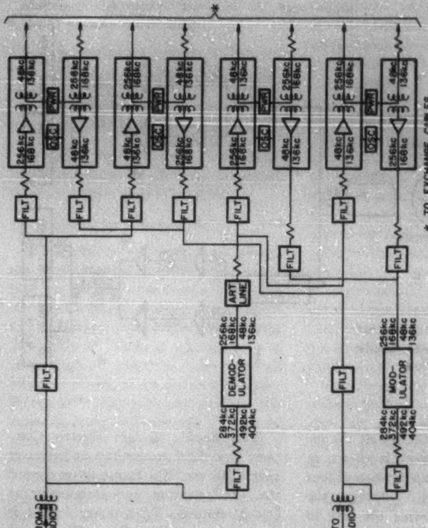
AT LOS ANGELES

Fig. 2 shows the arrangement for combining the four *N1* carrier groups at Los Angeles. The availability of the 12-channel *N1* equipment in two frequency allocations, 44 to 140 kc and 164 to 260 kc, made it possible to combine two 12-channel groups by means of band-pass filters; however, a means was needed for translating the remaining two systems to frequencies above 260 kc to enable the four systems to be combined for transmission over the radio facility. Since an *N1* repeater consists essentially of a modulator element in conjunction with an amplifier, the possibility of modifying a repeater to perform the new requirement for frequency translation was considered. If such a modification were feasible, it appeared that the method would have several advantages such as the following:



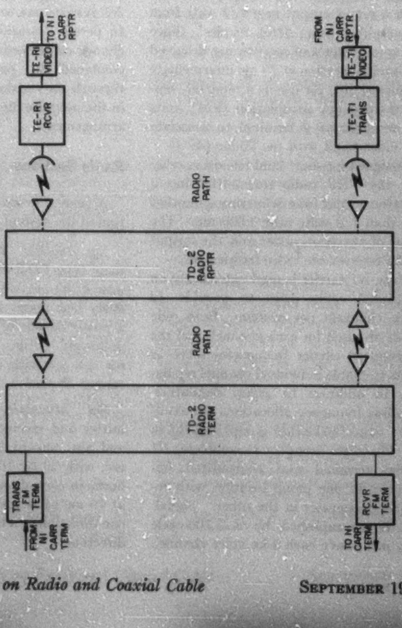
CHAN UNIT
TO 11 OTHER CHAN UNITS
HGT
LGR
LGT
HGR
256kc 168kc
48kc 136kc
ART LINE
FILT
MOD-ULATOR
DEMOD-ULATOR
284kc 372kc 492kc 404kc
48 CHAN TO RADIO
48 CHAN FROM RADIO

Fig. 2. N1 carrier terminals block diagram



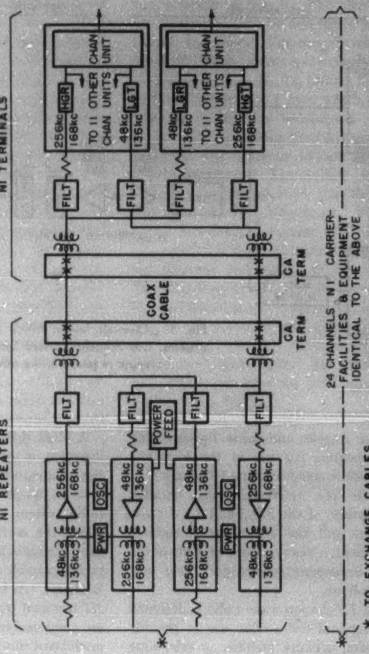
FROM RADIO
TO RADIO
FILT
DEMOD-ULATOR
MOD-ULATOR
ART LINE
284kc 372kc 492kc 404kc
256kc 168kc
48kc 136kc
OSC
PWR
*
* TO EXCHANGE CABLES

Fig. 4 (right). N1 carrier repeaters block diagram



TRANS FM TERM
FROM N1 CARR TERM
TD-2 RADIO TERM
RADIO PATH
TD-2 RADIO RPTR
RADIO PATH
TE-R1 RCVR
TE-R1 VIDEO
TO N1 CARR RPTR
RCVR FM TERM
TO N1 CARR TERM
TE-T1 TRANS
TE-T1 VIDEO
FROM N1 CARR RPTR

Fig. 3. Radio facility block diagram



N1 REPEATERS
N1 TERMINALS
48kc 136kc
256kc 168kc
OSC
PWR
POWER FEED
FILT
CA TERM
COAX CABLE
CHAN UNIT
TO 11 OTHER CHAN UNITS
HGR
LGT
LGR
HGT
24 CHANNELS N1 CARRIER-FACILITIES & EQUIPMENT IDENTICAL TO THE ABOVE
* TO EXCHANGE CABLES

Fig. 5. N1 carrier repeaters and terminals block diagram

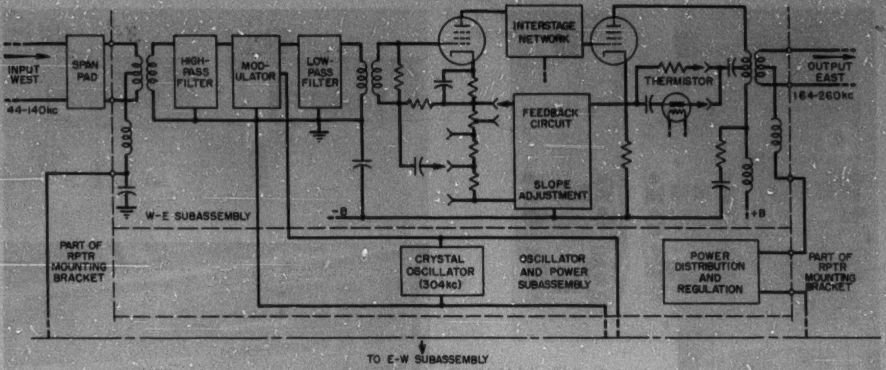

INPUT WEST
44-140kc
SPAN PAD
HIGH-PASS FILTER
MOD-ULATOR
LOW-PASS FILTER
INTERSTAGE NETWORK
THERMISTOR
FEEDBACK CIRCUIT
SLOPE ADJUSTMENT
OUTPUT EAST
164-260kc
W-E SUBASSEMBLY
-B
+B
PART OF RPTR MOUNTING BRACKET
CRYSTAL OSCILLATOR (304kc)
OSCILLATOR AND POWER SUBASSEMBLY
POWER DISTRIBUTION AND REGULATION
PART OF RPTR MOUNTING BRACKET
TO E-W SUBASSEMBLY

Fig. 6. N1 carrier low-high repeater

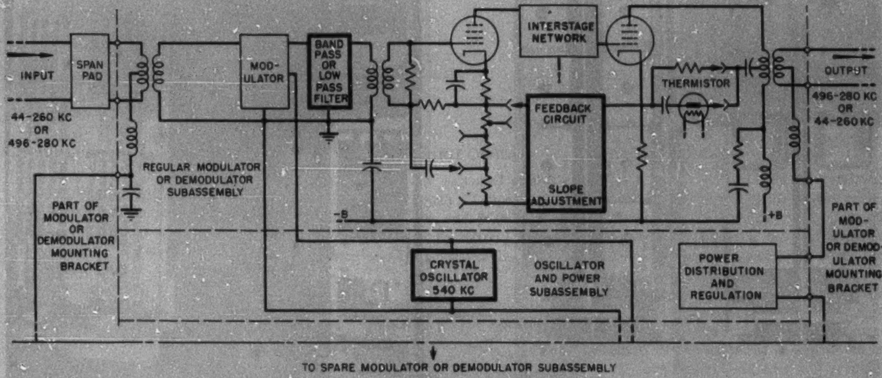

INPUT
44-260 KC OR 496-280 KC
SPAN PAD
MOD-ULATOR
BAND PASS OR LOW PASS FILTER
INTERSTAGE NETWORK
THERMISTOR
FEEDBACK CIRCUIT
SLOPE ADJUSTMENT
OUTPUT
496-280 KC OR 44-260 KC
REGULAR MODULATOR OR DEMODULATOR SUBASSEMBLY
-B
+B
PART OF MODULATOR OR DEMODULATOR MOUNTING BRACKET
CRYSTAL OSCILLATOR 540 KC
OSCILLATOR AND POWER SUBASSEMBLY
POWER DISTRIBUTION AND REGULATION
PART OF MOD-ULATOR OR DEMOD-ULATOR MOUNTING BRACKET
TO SPARE MODULATOR OR DEMODULATOR SUBASSEMBLY

Fig. 7. Modulator or demodulator

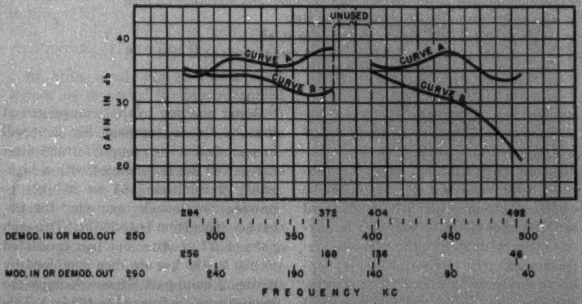


Fig. 8. Modulator-demodulator gain frequency characteristics. Curve A for modulator, curve B for demodulator

1. Production problems would be minimized.

2. Mounting arrangements, external wiring, and power supply would be identical to those of the unmodified repeater.

3. Maintenance routines and test equipment would be the same as for a repeater.

An *N1* low-high repeater, Fig. 6, was altered to perform the desired modulating function. Fig. 7 shows the circuit after modification. Major changes were as follows:

1. The carrier oscillator frequency was changed from 304 to 540 kc by replacing the crystal.

2. The slope circuit in the feedback path of the amplifier was disconnected to provide more nearly flat gain.

3. The input filter was removed and the output filter of the modulator was replaced by a band-pass filter to pass 280 to 496 kc,

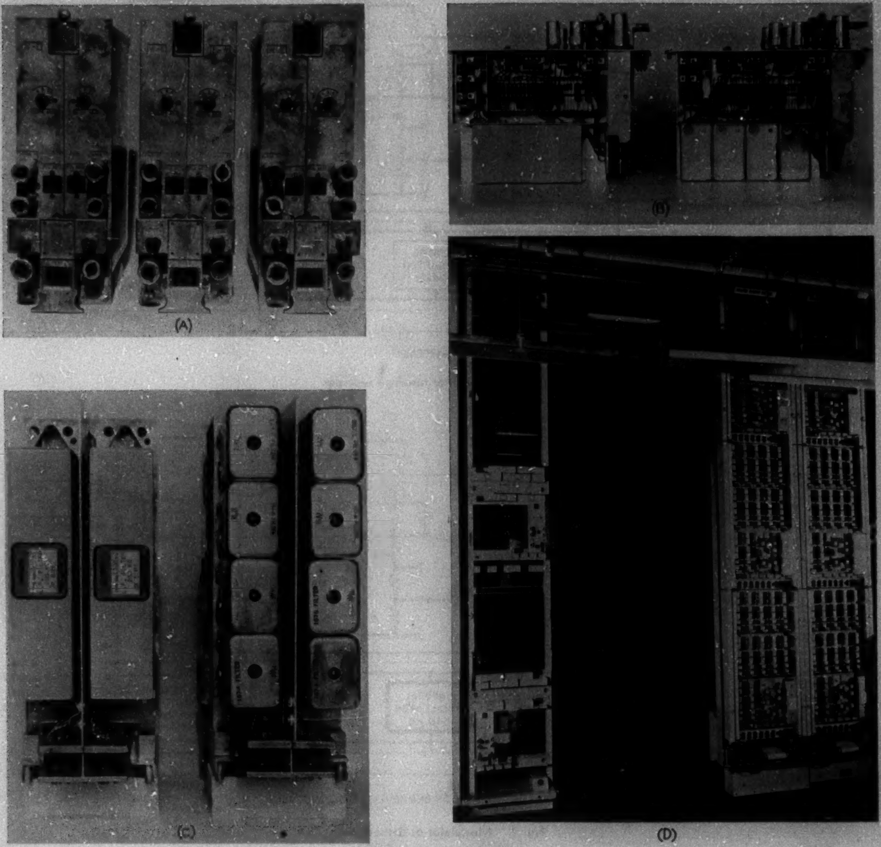


Fig. 9. The N1 carrier

A—Front view of N1 carrier low-high repeater at left, demodulator in center, and modulator on right
B—Side view of modulator at left and N1 carrier low-high repeater at right
C—Bottom view of N1 carrier low-high repeater at right and demodulator at left
D—N1 terminal at Los Angeles. Modulator and demodulator at left; N1 carrier terminals at right

the *A-B* products resulting from an input of 44 to 260 kc.

The demodulator is identical to the modulator except that a low-pass filter is used to pass the *A-B* products (44 to 260 kc) obtained with an input of 496 to 280 kc.

Fig. 8 shows graphically the gain-versus-frequency characteristics of the modulator and demodulator units. The variations in gain with frequency are largely attributable to losses in the transformers, which were originally designed for low loss at normal *N1* carrier frequencies. Replacement of transformers for this special application was not believed to be warranted, since: 1. constant slope variations in gain could be corrected through the use of standard artificial lines design for use with *N1* carrier along with slope adjustments in repeaters and receiving terminals; 2. nonlinear variations in gain, if not too large, could be cared for by individual channel regulation at the carrier-receiving terminals.

Again referring to the arrangement at the Los Angeles terminal, Fig. 2, it will be seen that a low-group *N1* transmitter (44 to 140 kc) is combined with a high-group transmitter (164 to 260 kc) to provide 24 channels occupying the frequency space from 44 to 260 kc. Two such pairs of *N1* systems (one pair translated to 280 to 496 kc) are then combined by means of band-pass filters to occupy the frequency range from 44 to 496 kc. This latter band of frequencies is then applied to the *TD-2* radio transmitter.

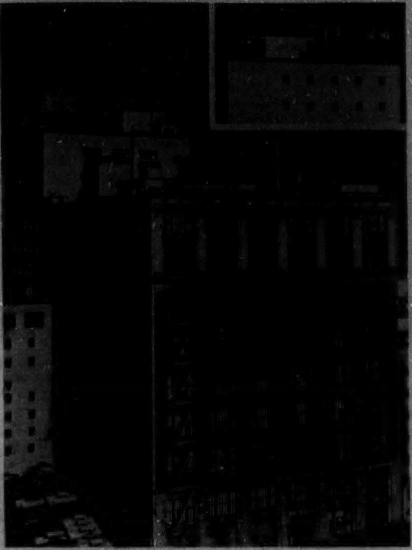


Fig. 10 (left). TD-2 antennas on the roof at Los Angeles

Fig. 12 (right). TD-2 radio terminal at Los Angeles

To obtain a reasonably flat input to the radio equipment, slope adjustments are made by the use of *N1* carrier artificial lines while level adjustments are obtained by using *N1* carrier span pads; further, the span pads, placed between the artificial lines and the filters, provide impedance matching at the filter inputs better than would be given by direct connection between the filters and artificial lines.

Fig. 9 (A) shows the modulator and demodulator alongside an unmodified *N1* carrier repeater. "Before" and "after" pictures of the units are shown in Figs. 9(B) and 9(C). The *N1* terminal arrangement at Los Angeles is pictured in Figs. 9(D).

AT VENTURA (MILLS ROAD)

At Mills Road, Fig. 4, the received signals are separated (by filters identical to those at the Los Angeles terminal) into two frequency groups: 44 to 260 kc and 280 to 496 kc. A demodulator, as previously described, is used to translate the 280 to 496-kc group to 44 to 260 kc. Each of the two pairs of *N1* frequency groups, 44 to 260 kc, is then filtered into a low group, 44 to 140 kc, and a high group, 164 to 260 kc. After passing through the usual *N1* repeaters at Mills Road, the groups are conducted individually on exchange cable pairs to the Ventura toll central office.

AT VENTURA CENTRAL OFFICE

At Ventura the 4-channel groups are again passed through *N1* repeaters, then combined in low- and high-group frequency pairs (44 to 260 kc) for transmission over the coaxial cable.

Fig. 11. TE-1 radio installation at Mills Road. Note alsynite window in ceiling

AT SANTA BARBARA

At Santa Barbara, the *N1* receiving terminal at the end of the coaxial cable, the groups of frequencies were separated by means of the same types of filters as were used at other locations. Here the individual channel regulation cares for any mop-up equalization which may be required.

In engineering the physical arrangement of the equipment in central offices two precautions were kept well in mind. First, since *N1* receiving frequencies identical to transmitting frequencies of other *N1* systems were used, care was taken to avoid excessive coupling between such systems. This was partially accomplished by providing separation in certain portions of the interconnecting cables and by separating transmitting and receiving filter circuits. Second, in the filter circuits and in the unbalanced circuit elements connected to them, grounding was at a single, common point to prevent stray currents.[8]

Test Results

Noise measurements were made under severe conditions of loading. With 1,000-cycle tones applied at Los Angeles on all channels except the one being measured, the average single-channel noise for the 48 channels as measured at Santa Barbara was 21.4 decibels adjusted

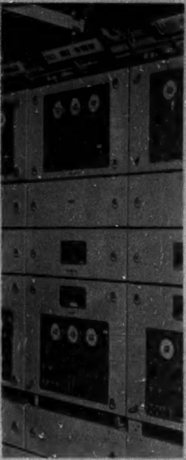


Fig. 13 (left). TD-2 70-mc FM terminal at Los Angeles

Fig. 14 (right). Plane antenna reflector at Mills Road

(dba), the maximum was 30 and the minimum was 14. All measurements were at a −9-decibel level point referred to the transmitting switchboard. With all channels idle the average single-channel noise was less than 8 dba, the maximum 17, and the minimum less than 5.

Tests were made to indicate the performance of the *N1* carrier signalling, which uses a 3,700-cycle tone above voice on each channel. These tests revealed that no signalling difficulties were to be expected in this application; however, the possibility was indicated that during deep radio fades the false seizure of senders in an intertoll switching arrangement could be a problem. Further investigation would be required if such usage were planned.

Maintenance

The entire layout was designed so that no exceptional maintenance equipment or procedures would be required. Major equipment items are in common use with standardized instructions available. Modulator and demodulator units are so closely similar to *N1* repeaters that special treatment is at a minimum. Filters, which were the only other units developed for the system require only the low degree of maintenance associated with inactive elements. Performance of the carrier systems through a period of more than a year has shown that reliable service with good transmission can be expected from a system such as the one described.

References

1. The TD-2 Microwave Radio Relay System, A. A. Roetken, K. D. Smith, R. W. Friis. *Bell System Technical Journal*, New York, N. Y., vol. 30, 1951, pp. 1041–77.

2. An Unattended Broad-band Microwave Repeater for the TD-2 Radio Relay System, R. W. Friis, K. D. Smith. *Electrical Engineering*, vol. 70, Nov. 1951, pp. 976–81.

3. Frequency-Modulation Terminal Equipment for the Transcontinental Relay System, J. G. Chaffee, J. B. Maggio. *Ibid.*, Oct., pp. 880–83.

4. A New Microwave Television System J. F. Wentz, K. D. Smith. *AIEE Transactions*, vol. 66, 1947, pp. 465–70.

5. A New Telephone Carrier System for Medium-Haul Circuits, R. S. Caruthers, H. R. Huntley, W. E. Kahl, Ludwig Pedersen. *Electrical Engineering*, vol. 70, Aug. 1951, pp. 692–97.

6. The Type N1 Carrier Telephone System: Objectives and Transmission Features, R. S. Caruthers. *Bell System Technical Journal*, New York, N. Y., vol. 30, 1951, pp. 1–32.

◆

No Discussion

Comparative Propagation Studies on 250 and 450 Megacycles

EMIL HOPNER
NONMEMBER AIEE

T. D. CUSHING
NONMEMBER AIEE

THE existence of two identical over-water radio links, one in the 250- and one in the 450-megacycle (mc) range, provided an opportunity for comparative studies of propagation in the two frequency bands. In making the studies, received signal strengths were recorded and shown as: 1. signal strength as a percentage of time, and 2. as signal strength probability. Observed signal strengths were compared to calculations made on the basis of previously published theoretical data.

An excellent opportunity for comparative propagation studies on 250 and 450 mc was provided by two existing over-water links operated by the Northwest Telephone Company in British Columbia. The two links provide a total of 16 voice channels between the city of Parksville on Vancouver Island and the city of Vancouver on the mainland. Since the terminal equipment and antennas for both links are in the same locations at both Vancouver and Parksville, identical propagation conditions exist. Propagation comparisons were made by calculating the expected signal strengths at the two receivers and comparing these calculated values with the actual values received over periods of 2 and 3 weeks.

A profile of the Vancouver-Parksville links is shown in Fig. 1. Path length is 52.5 miles. Effective antenna heights are 275 feet at Parksville and 117 feet at Vancouver. The two receivers at Vancouver are shown in Fig. 2, and the antennas at Vancouver are shown in Fig. 3.

Signal Strength Calculations

The received signal strength to be expected on the two links was determined on the basis of methods of calculation previously published by Bullington.[1] The actual link frequencies used were 243 and 455.25 mc. Transmitter output powers were 25 watts [+14 decibels above 1 watt (dbw)] and 50 watts (+17 dbw)

Paper 54-241, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 22, 1954; made available for printing May 3, 1954.

Emil Hopner is with the Lenkurt Electric Company, Inc., San Carlos, Calif., and T. D. Cushing is with the Northwest Telephone Company, Vancouver, B. C., Canada.

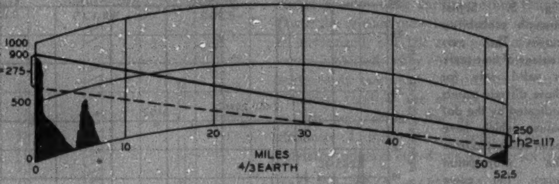


Fig. 1. Path loss figures and profile for the Vancouver-Parksville radio links used in these propagation studies. Path length 52.5 miles bearing 91 degrees true

Path Loss	234 Mc, Db	455.25 Mc, Db
Calculated, db		
Plane earth, $h_1=275$ feet		
$h_2=117$ feet		
$d=52.5$ miles	123	123
Measured, db		
90% value	127	123
Most probable path loss	124	122
Fading below 90% value	−11	−22

respectively; antenna gains were 12 and 15 decibels (db) respectively; and coaxial cable losses for each equipment terminal were approximately 1.5 db.

Thus, total gain for the 243-mc link was:

Antenna gain, db	+24
Coaxial cable loss, db	−3
Power output, dbw	+14
Total gain, db	+35

For the 455.25-mc link, total gain was:

Antenna gain, db	+30
Coaxial cable loss, db	−3
Power output, dbw	+17
Total gain, db	+44

The expected field strength for links of this type is the sum of the directed and reflected waves and is calculated from the equation

$$E=\frac{2E_0}{d}\sin\left(\frac{2\pi h_1 h_2}{\lambda d}\right)$$

where

E = field strength at receiver
E_0 = strength of direct wave at unit distance
d = distance between antennas
λ = wave length, same units as d
$h_1 h_2$ = effective heights of sending and receiving antennas, same units as d

The term E_0/d represents the loss due to the direct wave (free space loss). For the 243-mc link E_0/d equals 114 db, and for the 455.25-mc link E_0/d equals 119.5 db. The term $2\sin[(2\pi h_1 h_2)/(\lambda d)]$ represents the loss or gain introduced by the reflected wave. This loss is about 9 db for the 243-mc link and about 3.5 db for the 455.25-mc link. Thus, total path loss for both links is approximately 123 db (114+9 and 119.5+3.5). Consequently, under these conditions the total path loss is not a function of frequency.

Expected signal strength for the two radio links is the sum of the total path loss and the system gain. For the 243-mc link this is:

Path loss, db.......−123
System gain, db.... +35
Expected signal strength =..... −88 dbw≅ 280 microvolts (mv)

For the 455.25-mc link:

Path loss, db..........−123
System gain, db....... +44
Expected signal strength =........ −79 dbw≅800 mv

Received Signal Strength Measurements

The strengths of the signals received from Parksville were recorded at Vancouver during the month of June 1953. Signals were recorded for 2 weeks on the 243-mc link and for 3 weeks on the 455.25 mc link. Results of the recorded measurements are shown in Figs. 4 and 5 plotted as received signal strength probability and received signal strength in percentage of time. Minimum received signal strength was about the same for both systems: 50 mv for the lower frequency, 60 mv for the higher. Maximum received signal strength was 1,350 mv at 243 mc and 3,500 mv for 455.25 mc.

As shown in Fig. 4, the received signal strength at 243 mc was above 175 mv, or 4 db below the calculated expected value, for 90 per cent of the total recording time. On the 455.25-mc link, also shown in Fig. 4, the received signal strength was above 770 mv, or only 0.3 db below the calculated expected value, for 90 per cent of the time.

Both of the signal strength probability

Fig. 2 (left). Receiver room at William Farrell Building, Vancouver. The 455.25- and 243-mc receivers from Parksville are at lower right and middle left respectively

Fig. 3 (right). Antennas on roof of William Farrell Building, Vancouver. The 455.25- and 243-mc antennas are in foreground and background respectively

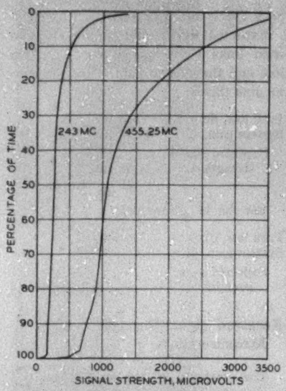


Fig. 4. Signal strength as a percentage of time for both Vancouver and Parksville links. Total recording time: 29,203 minutes for 455.25 mc; 21,655 minutes for 243 mc

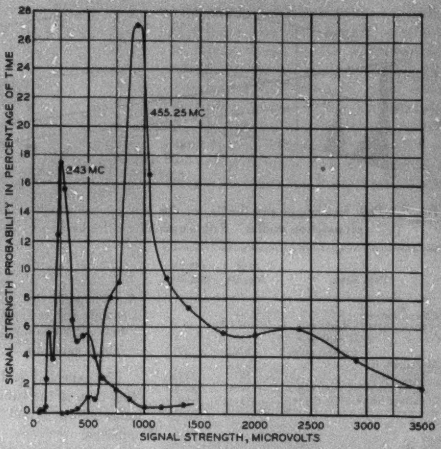


Fig. 5. Signal strength probability curves. The percentage of time scale is valid only for those discrete points indicated by the dot values which have been computed from the propagation graph. The curve formed by joining these discrete points shows only a relative indication of the amount of time any particular signal strength value might be expected. Total recording time is the same as in Fig. 4

curves shown in Fig. 5 have a very pronounced maximum value. This maximum is within $+1$ db of the expected value for both the 455.25- and the 243-mc links. It is interesting to note that the maximum probability value is much higher for the 455.25-mc link than for the 243-mc link.

The total fading range encountered during the tests was 29 db for the lower frequency link and 35 db for the higher. However, the fading range below the 90-per-cent value is 11 db for 243 mc and 22 db for 455.25 mc. Consequently, it appears that necessary fade margins for overwater paths such as these are approximately 0.25 db per mile in the 250-mc range and approximately 0.5 db per mile in the 500-mc range.

These results are within the measuring accuracy of the tests. It should be pointed out that the stronger signals on 450 mc are attributable to increased power (3 db), increased antenna gains (total 6 db), and first fresnel zone path clearance. Because of the increased noise figure for the 450-mc receiver, about 4 db stronger signal would be required to give the same resulting signal-to-noise ratio. While these recordings cover a period of about 3 weeks, the 243-mc circuit has been in operation 3 years and the 455.25-mc circuit over $1^1/_2$ years, and the performance over this period agrees with this report.

It would be interesting to carry out tests on 900 mc on a similar overwater path if sufficient antenna gains could be obtained to overcome the lower power, greater fading range, and increased noise figure.

References

1. Radio Propagation at Frequencies Above 30 Megacycles, Kenneth Bullington. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 35, Oct. 1947, pp. 1122–36.

◆

No Discussion

Thickness Gauge for Dielectric Materials

W. W. WOODS
NONMEMBER AIEE

THE instrumentation described in this paper is based upon the variation of mutual inductance between two coils when brought near a metal surface. A type of thickness gauge employing this principle was first publicized by M. L. Greenough[1–4]; he stated that the signal obtained from a mutual inductance was "greater and more linear" than from self-inductance. Since then applications of the principle have been extended to the measurement of distance and vibration.

The principal advantage of the use of a mutual inductance change for the measurement of displacement lies in the fact that distances to a metal surface may be measured without an application of appreciable forces. Any dielectric material—solid, liquid, or gaseous—may be interposed. This enables thickness measurements on large dielectric fabrications and displacement measurements on delicate systems with equal facility.

Basic Operation

A simple representation of the principle of operation of the mutual inductance element can be made by employing the concept of images. When a coil of wire carrying an alternating current is brought near a perfectly conducting plane, the resulting field is identical to that which would be produced by the same coil and its image in the absence of the plane. This follows directly from electromagnetic theory.[5] Thus, as shown in Fig. 1, the mutual inductance between two coils A and B in the presence of a conducting plane may be said to be the mutual inductance M_{AB} between the two coils in free space less the mutual inductance

---

Paper 54-296, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 26, 1954; made available for printing May 24, 1954.

W. W. Woods is with the Boeing Airplane Company, Seattle, Wash.

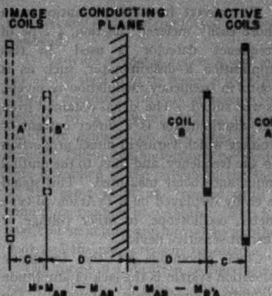


Fig. 1. Image coil concept

$M_{AB'}$ between one coil and the image of the other. In usual practice, the two coils are mounted coaxially on a solid dielectric form, with the separation C of the coils fixed. Changes in the distance D from the plane to the coil system cause a variation of the image mutual inductance $M_{AB'}$.

The curves of mutual inductance as a function of the distance D from the conducting plane can be computed for simple systems[6,7] similar to those shown in Fig. 1. An idea as to the general shape of such curves may be gained from Fig. 2. Here are represented nondimensional curves of mutual inductance versus distance for several geometries of multiturn coils. Calculations were made on the basis of coils of square cross section, each side of this square being 1/40th of the diameter A of the larger coil. It is to be seen that a relatively linear portion of the curve occurs about an inflection point and covers a range of about 7 per cent of the larger coil diameter. For vibration-measuring instruments, and for measuring small distances, this is the portion of the curve usually used. However, for thickness measurements of large dimension, the full curve is usable.

In practice the reflecting surface used need be neither infinite in extent nor of infinite conductivity. The practical limit of observable effects upon readings is reached for reflector dimensions greater than twice the coil system diameter and a thickness exceeding three times the electrical depth of penetration, or skin depth, for the frequency and metal used. This skin depth should be small with respect to the distance to be measured. For special purposes any or all these limits may be violated, resulting in systems wherein calculated responses must be modified by empirical data.

The foregoing applies to plane-reflector surfaces only. For surfaces other than plane the mutual inductance curve shape is unchanged, but the zero point of the curve is shifted. For circular cylindrical surfaces an empirical equation has been obtained relating the system geometry to the zero error. For an outer coil diameter A, inner coil diameter B, and reflector radius of curvature R, the error is

$$e = (1.5A^2 + 5.8AB - 2.4B^2)/100R \qquad (1)$$

To obtain this equation, a system of nine coils was fabricated. For each 2-coil combination values of mutual inductance were obtained with the coil form pressed against conducting metal cylinders of various radii of curvature. These data were reduced to give curves of effective displacement referred to a flat surface versus radius of curvature of the cylinder. Standard curve-fitting techniques were used to obtain equation 1 from this mass of data. Subsequent checks using dielectric spacers of known thickness between the coil form and the cylinders gave data which complied with equation 1 within experimental error, thus verifying the invariance of the curvature correction with displacement.

For surfaces other than plane or cylindrical, no data have been tabulated to date which can be used for accurate prediction of error. Where a compound surface has two radii of curvature, one much larger than the other as in an ogive, equation 1 for cylindrical surfaces will hold to a close approximation using the smaller radius of curvature.

Various dielectrics appear to have little effect upon this type of instrument. A small effect is noted for large coil systems which appears, to a first approximation, to be a zero change proportional to the frequency of operation and thickness and dielectric constant of the dielectric in the field of the instrument.

Instrumentation Methods

To render a mutual inductance device useful for indicating distances or displacements, a method for measuring mutual inductance must be provided. Several instrumentation methods are applicable using alternating currents ranging from audio to video frequencies. These methods may be divided into three general classes: passive bridge networks, voltage amplitude metering systems, and frequency deviation systems.

Standard bridge networks for the measurement of mutual inductance, including impure mutual inductance, can be found in texts on electrical measurements. The more common of the bridge networks are ascribed to Hartshorn and Carey Foster. Two disadvantages of this type of measurement are the relatively tedious manual operation and the necessity for very close connections to the mutual inductance being measured. The use of long leads or cables to the coil system make measurement difficult, if not impossible.

For the convenience of direct indication and to allow the use of moderate cable lengths, a voltage amplitude system may be used. Voltage amplitude metering systems take advantage of the fact that the amplitude of the voltage induced in one coil by an alternating current in the

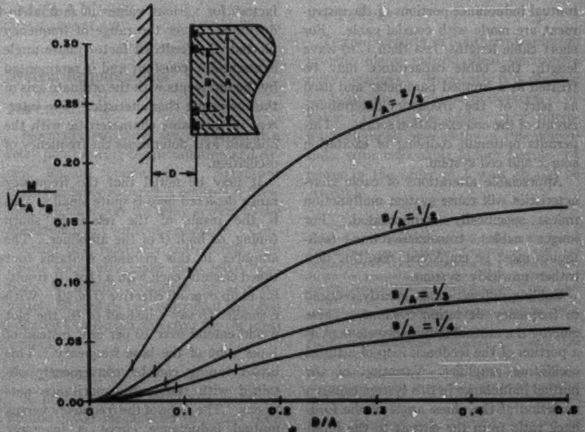


Fig. 2. Mutual inductance versus separation

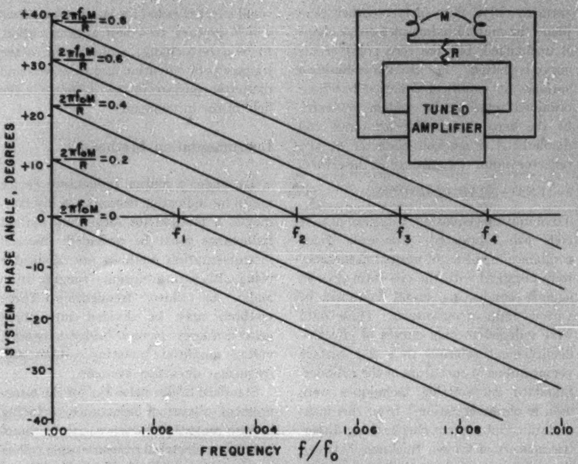


Fig. 3. Frequency deviation system

other is equal to the product of the mutual inductance, the angular frequency, and the current amplitude. For the induced voltage to bear a constant proportionality to the mutual inductance, the product of the frequency and the current amplitude must be held rigidly fixed. This may be accomplished by holding each factor constant, or by allowing one to vary and regulating the other to produce a constant product. Both methods have been successful in instrument applications.

Connections between the metering and mutual inductance portions of the instrument are made with coaxial cable. For short cable lengths, less than 1/10 wave length, the cable capacitance may be treated as a lumped parameter and used as part of the frequency-determining circuit of the coil excitation source. This permits optimum coupling of excitation source and coil system.

Appreciable alterations of cable characteristics will cause system malfunction unless specifically compensated. For longer cables transmission-line techniques must be employed, resulting in a rather unwieldy system.

Cabling difficulties are greatly reduced in frequency deviation systems. These employ the mutual inductance element as a portion of the feedback loop of a tuned oscillating amplifier. Variation of the mutual inductance in this system causes a variation of the phase angle of the feedback ratio from the output to the input of the amplifier portion of the oscillator. Since the total loop phase angle must equal zero—or an integral multiple of 360 degrees—for oscillations to continue, the frequency of oscillation will change, altering the phase angle of the amplifier gain factor to compensate the change in the feedback ratio. The frequency of oscillation then is a measure of the mutual inductance in the feedback loop. An illustration of this is shown in Fig. 3.

Here are shown the phase characteristics of a system presently in use. The family of curves represents the combined phase angles of the gain and feedback factors for various values of mutual inductance. Over the range of frequency plotted, the feedback factor phase angle is essentially constant and is represented by the intercepts with the ordinate axis of the composite characteristic. The intercept of this same characteristic with the abscissa axis determines the frequency of oscillation.

It may be noted that the frequency range depicted here is quite small. This is the result of the relatively sharp tuning, or high *Q* of the amplifier. The amplifier in this instance contains four tuned circuits, each with a *Q* of 10, resulting in an over-all effective *Q* of 40. With a smaller *Q*, say 4 instead of 40, the plot would extend over 10 per cent instead of 1 per cent of the base frequency. This broad tuning can be conveniently obtained with resistance-capacitance networks. The slope of the frequency versus mutual inductance curve is inversely proportional to the *Q* of the system.

To convert frequency variation to a more easily metered function, a type of frequency detector is used. In this application a discriminator, such as is found in frequency-modulation receivers, is well suited. The signal obtained from the discriminator is a direct current or voltage which varies in direct proportion to the frequency, and thus, to the mutual inductance being measured. This signal is easily displayed on a D'Arsonval-type meter, oscilloscope, or other voltage or current-sensitive device.

Chief advantage of the frequency deviation system is the lack of amplitude stability requirements throughout the main portion of the system. Phase stability is the criterion here, and is in general a bit easier to maintain. Cable lengths up to several wave lengths may be employed with little difficulty. Zero adjustments are easily made on this system by using stable receiving-type tuning elements in convenient portions of the circuit. The frequency response of the system is dependent primarily on the frequency metering circuit, which is capable of wide range adjustment. Additional frequency meters may be provided at remote locations with little difficulty.

The frequency deviation system is readily adapted to a mechanical servo type of indicator or recorder. The output of the discriminator provides the signal for controlling a motor driving phase adjusting elements in either the amplifier or mutual inductance circuits. An indicator or recorder can be coupled to the motor shaft.

An example of an application of the foregoing discussion is shown in Fig. 4. In this case the mechanical thickness of a large dielectric shell was desired. Owing to the size and shape of the shell, use of mechanical micrometers was inconvenient and difficult. A means of measurement was needed both while the shell was undergoing manufacture on a cast-iron

Fig. 4. Thickness gauge in operation

mandrel and after the completed part was removed.

To provide a metal reflector on one side of the shell after its removal from the forming mandrel, a piece of aluminum foil backed by a resilient pad is pressed manually against the shell from its interior. The mutual inductance coil structure is applied to the outside of the shell at the same point and rocked until a minimum reading is obtained. The foil and coil structure are moved to another spot and the measurement repeated.

The range of indication of this instrument is 0.2 to 0.4 inch. The indicated reading must, of course, be corrected for curvature of the shell at the point being measured. As shown in Fig. 4, the surface is not cylindrical but has a continuously varying curvature. A correction curve was established by finding the indication error at various points on the mandrel and plotting this as a function of distance along the mandrel axis. The appropriate correction for any point on the dielectric surface can be determined from this curve and subtracted from the reading to obtain actual thickness. Measurements so obtained are reproducible within less than 1 per cent.

The mutual inductance sensing element of this instrument was constructed with coil diameters of 0.5 and 2 inches. This configuration was chosen as a compromise between linearity, curvature error, and sensitivity.[8]

Other applications of the mutual inductance principle are currently being made to problems of vibration and coating thickness measurement. An instrument employing the frequency deviation system is now being employed to measure vibration amplitudes of the order of a few microinches peak amplitude at frequencies of the order of kilocycles in the presence of much larger amplitude low-frequency vibration.

In conclusion it may be stated that an instrumentation method has been developed to the point of predictable design for the measurement of distance and vibration amplitude through the distortion of a magnetic field by conductors. Practical instrumentation may be designed for full scale ranges as low as 0.005 inch and as high as 1 inch. Measurement may be made through a wide variety of dielectrics to an otherwise inaccessible metal surface, and the absence of appreciable forces inherent in the measuring process allows use on delicate systems. High stability, rapid response, and remote indication are obtainable without great cost or complexity.

References

1. Technical Data on Electronic Micrometer, M. L. Greenough. *Electronics*, New York, N. Y., vol. 20, no. 11, Nov. 1947, p. 172.

2. Electronic Micrometer, M. L. Greenough. *Radio News*, Chicago, Ill., Radio Electronic Engineering Edition, Aug. 1947.

3. Oil Film Thickness Indicator for Journal Bearings, M. L. Greenough. *AIEE Transactions*, vol. 67, pt. I, Aug. 1948, pp. 589–95.

4. An Electronic Circuit for Measuring the Displacement of Pressure-Sensitive Diaphragms, M. L. Greenough, W. R. Williams. *Journal of Research*, National Bureau of Standards, Washington, D. C., vol. 46, no. 1, Jan. 1951.

5. Ultra High Frequency Techniques (book), J. G. Brainer, G. Koehler, H. J. Reich, L. F. Woodruff. D. Van Nostrand and Company, Inc., New York, N. Y., 1942, p. 206.

6. Fields and Waves in Modern Radio (book), S. Ramo, J. R. Whinnery. John Wiley and Sons, Inc., New York, N. Y., 1944, pp. 224–25.

7. Radio Engineers' Handbook (book), F. E. Terman. McGraw-Hill Publishing Company, Inc., New York, N. Y., 1943, pp. 67–70.

8. A Mutual Inductance Thickness Gage, W. W. Woods. *Document 13851*, Boeing Airplane Company, Seattle, Wash., April 1953.

No Discussion

One Approach to a Video SHF Relay System

R. H. COE
NONMEMBER AIEE

F. F. McCLATCHIE
ASSOCIATE MEMBER AIEE

THE authors' company is engaged in common carrier communications service in Southern California. One of the many services offered by the company is the provision of video transmission facilities between any points specified within the area. Permanent transmission facilities have been provided over principal demand (backbone) routes and portable microwave relay systems interconnect branch points to the permanent system.

An ideal microwave relay system for this service application must be readily portable by the personnel assigned to operate relay points, designed to operate in a portion of the frequency spectrum allocated for common carrier use, and have available sufficient radio-frequency (r-f) power to meet the path length requirements. These requirements will be individually considered.

System Requirements

Portability

Normally two men are available at a relay point during the installation and circuit line-up operation. A significant saving in labor expense is realized if these men can transport the relay system to the operation point without assistance. Considering that relatively inaccessible high points frequently make ideal relay locations, a weight limitation of approximately 50 pounds per component has been adopted as a design objective for the ideal system.

Frequency

A portion of the frequency spectrum around 6,000 megacycles (mc) has been allocated for common carrier use. There are reliable circuit components commercially available which may be incorporated in the design of a system for operation in this region. It therefore is taken as a design objective that the ideal relay system will utilize frequencies in the 6,000-mc band.

R-F Power

Terrain features in Southern California provide many natural ideal microwave relay locations. Frequently it is possible to obtain line of sight paths in excess of 50 miles. To provide high-quality microwave channels over paths of this magnitude, with reasonable fading margin, r-f power in the order of 1 watt should be available.

Paper 54-239, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 2, 1954; made available for printing April 27, 1954.

R. H. Coe and F. F. McClatchie are with the Pacific Telephone and Telegraph Company, Los Angeles, Calif.

For information furnished in connection with the preparation of this paper the authors are indebted to F. W. Bailey of Lambda Pacific Engineering for data on wave-guide components, J. H. Clark of Pacific Telephone and Telegraph Company for data on color transmission, and E. M. Hitchcock of Pacific Telephone and Telegraph Company for data on interference suppression.

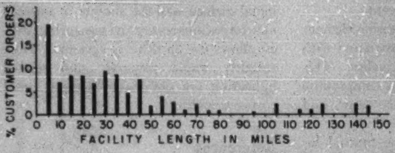


Fig. 1. Customer orders versus total length of facility. Based on 242 orders for service between 9-1-1949 and 1-20-1954

Choice of System

A study of the economic aspects of the ideal system versus available microwave relay systems would in part consider total length of path per customer order, simultaneous service demand, relay systems required to meet customer requirements, and plant investment necessary to meet customer requirements. Figs. 1 through 4 tabulate this information. The total plant investment required to meet service demand is not used in these considerations. These figures represent only the costs of the microwave relay systems and no maintenance, testing, transportation, or other expenses have been included.

It was determined from considerations similar to the foregoing that the ideal system would be a desirable facility for use in the company's operations and steps were taken to procure such a system. It was found that none was available in sufficient quantities to meet the anticipated demand but that existing Radio Corporation of America (RCA) *TRR/TTR* systems could be modified to meet the requirements of the ideal system.

Many factors entered into the considerations prior to the choice of a system for operation in the 6,000-mc band. Two of the more pertinent were that most of the company operating personnel were familiar with RCA systems and that a significant number of the systems were in company plant. All factors being adequately considered, it was decided that system requirements could be met with modified RCA relay systems.

A brief technical summary of the electrical specifications of RCA *TTR/TRR* and RCA *TTR/TRR* modified *SHF* relay systems is presented in Table I for those not familiar with this microwave relay equipment.

Fig. 5 shows the complete relay system, less antennas, connected through a variable attenuator for testing.

General Modification Considerations

The modification incorporates in the basic 5-unit relay system the enunciated design criteria as well as improvements which were suggested as a result of experience gained from 462 links established under varied conditions since 1949. A discussion of these field-sponsored improvements follows.

An integral variable wavemeter, accessible to the operator during system use and illuminated for night operation, has been incorporated in the transmitter unit and receiver unit design. Fig. 6 shows the transmitter wavemeter in position. Minimal government requirements may be met with a fixed transmitter wavemeter but the method of operation which follows from this arrangement is extravagant in maintenance spare equipment and operator time required for system alignment. Variable wave meters provided at the transmitter and receiver permit the rapid interchange of one maintenance spare equipment with any other system in the event of relay link failure.

The wavemeter being accessible to the operator during system use permits the following additional applications: At the transmitter it is possible to monitor klystron frequency drift, execute a klystron linearity adjustment, test the klystron deviation sensitivity, and monitor the video signal contained in the r-f output of the transmitter; at the receiver location it is possible to measure transmitter frequency, rapidly realign the system after replacing the local oscillator tube, and identify the frequency of any interfering *SHF* signal.

A video input switching arrangement has been provided at the transmitter which selects one of two terminated customer feeds, an external test signal input, or an internally generated 60-cycle test level. This arrangement facilitates system alignment, in-service switching of customer video signals, and rapid isolation of trouble conditions.

A new transmitter video amplifier provides 32-decibel (dbv) gain. There are in out company plant *TE* relay systems which are nominally rated at 1-volt peak-to-peak video output. This is below the 1.4-volt peak-to-peak video input level nominally specified for the *TTR* relay

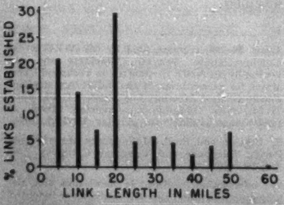


Fig. 2. Relay links established versus link length. Based on 460 relay links established between 9-1-1949 and 1-20-1954

Fig. 3 (right). System tandem operation limits. Relay systems versus total facility length

System A—2/10 watt, 57-inch antenna, 52-db video peak to peak/rms noise

System B—1/10 watt, 72-inch antenna, 52-db video peak to peak/rms noise, RCA TTR/TRR

System C—1 watt, 72-inch antenna, 52-db video peak to peak/rms noise, RCA TTR/TRR modified

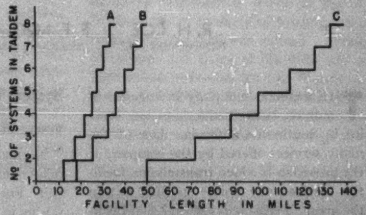


Fig. 4 (right). Based on Fig. 3 and cost of relay systems only

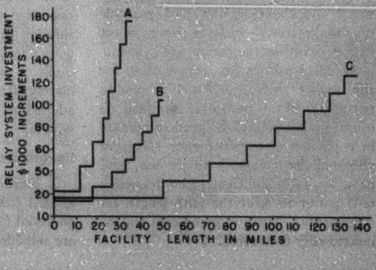


Table I. Electrical Specifications of RCA Relay Systems

	RCA TTR/TRR	RCA TTR/TRR Modified
Emission	frequency modulation	frequency modulation
Frequency band, mc	6,500–7,050	5,900–7,200
Transmitter frequency deviation each side of center, mc	6	4
Nominal video input, volts peak to peak	1.4	0.5
Receiver intermediate frequency (i-f), in mc, centered about	129	129
Receiver i-f and discriminator bandwidth, mc	16	20
Nominal power output of transmitter, watts	0.1	1
Nominal video output of receiver, volts peak to peak	1.4	1.5
Number of subunits	5	5

transmitter. The new nominal video input to transmitter level of 0.5 volt peak to peak made possible by the transmitter redesign eliminates a requirement for a portable amplifier, which was formerly interposed between the receiver of the *TE* relay systems and *TTR* transmitter units, when tandem operation was utilized.

Two operator communication systems have been provided which permit the use of either local battery or common battery telephone sets. Power for the common battery system is obtained from the filament supply through a suitable rectifier and filter combination. The local battery communication system may be interconnected to telephone company 2-wire maintenance circuits where regulations permit.

Meter facilities are provided for input voltage and all regulated potentials. The meter selector switch can be operated through the various steps without introducing picture interference and a transformer ratio selector switch has been installed to facilitate operation in areas where commercial power sources are not well regulated.

An interlock arrangement for personnel protection has been inserted in the system power supply leads. It is recognized that maintenance personnel must disable this protection arrangement when working on these units. The interlock disabling operation creates the possibility that systems might be returned to service with the interlock system electrically by-passed. A positive safeguard against this dangerous possible error is provided in a mechanical interlock by-pass that simultaneously prevents the positioning of the unit for service when personnel interlock protection is disabled. Fig. 7 shows the interlock by-pass in its on and off positions.

The susceptibility of the receiver to extraneous intermediate amplifier interference has been reduced by shielding the receiver unit, double shielding control, and camera cables and filtering all exposed leads. The combined result of these measures is better than 60-db suppression of a test interfering signal at 127.7 mc. Fig. 8 shows the grounding ring and finger stock arrangement applied to the receiver unit.

A differential phase and differential gain equalizer applique has been designed for use when *NTSC* color signals are transmitted over the system. In this connection three separate 2-link services have been established which measured zero differential phase and zero differential gain when properly equalized.

Considering each component of the relay system individually, the following steps are taken to modify the system.

Modification Procedure

Transmitter Unit

The transmitter unit is illustrated in Fig. 9. Choice of the wave-guide coupled *X26* klystron specified the design of a new wage-guide section, air cooling system, and mechanical tuning drive. The transmitting wave guide consists of a straight section of *RG-50-U* guide connecting the klystron to the antenna feed horn and a 30-db directional coupler for sampling the transmitted energy without affecting the operation of the klystron. An adjustable probe is located in the low level portion of the directional coupler to sample the transmitted energy. This probe is connected to the wavemeter assembly through a detachable coaxial cable. The resulting wave-guide assembly presents a standing wave ratio of not more than 1.05 to 1 to the klystron when the wave guide is properly terminated.

The proximity of the blower motor to the klystron presents some klystron modulation problems. The magnetic field surrounding the blower motor may frequency-modulate the klystron by its action upon the bunching of electrons in the electron beam. This modulation is reduced below the residual 60-cycle level by positioning the motor for minimum modulation.

The klystron mechanical tuning drive is geared so that full rotation of the shaft will cause a frequency change of approximately 25 mc. This is the frequency shift necessary to tune to an adjacent channel. As a further operating convenience, both the wavemeter and tuning control operate on the frequency in the same manner when rotated in the same direction.

A coaxial-type wavemeter was found

Fig. 5 (left). Relay system assembled for test

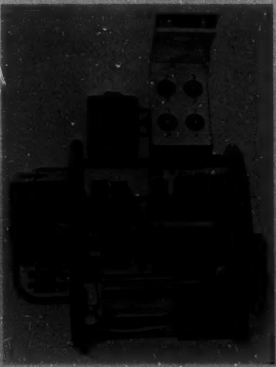


Fig. 6 (right). Transmitter wavemeter mounted in unit

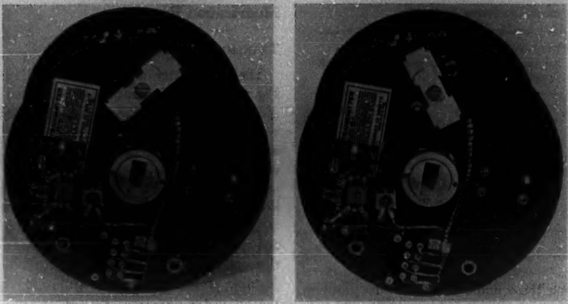


Fig. 7. Transmitter interlock by-pass arrangement in on and off positions

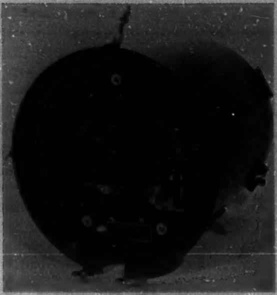


Fig. 9. Transmitter unit

to be the most durable and least expensive for use in the confines of the transmitter unit. The coaxial input to the wave meter is particularly well suited to the coaxial nature of the crystal matching network and flexible coaxial connection to the wave guide. Fig. 10 shows the wavemeter and crystal matching network. The coaxial input connects at point *A*. The 1/4 wave section is a mechanical support for the other sections and has little effect on the performance of the wavemeter over the 5.8 to 7.2-kilomegacycle operating range. The optimum crystal location is at point *A*. To circumvent mechanical difficulties, the crystal was located at the end of a half-wave section. This section transfers the impedance of the crystal to the end of the interconnecting cable. Another half-wave section transfers the impedance of the wavemeter to the end of the interconnecting cable. A reactance is placed in series with the half-wave section coupling to the wavemeter to decouple the wavemeter from the crystal and decrease loading of the wavemeter by the crystal. This reactance is adjusted until crystal current reduction caused by the resonance of the wavemeter is one-half of the off-frequency crystal current. This design makes posisible the mounting of a wavemeter at a point distant from the wave-guide plumbing. This distance is limited only by the loss in the coaxial cable, the sensitivity of the meter, and the available r-f power.

The wavemeter scale is illuminated by a dial light to facilitate system operation after dark. This light simplifies the work of field forces operating the equipment under adverse weather conditions in exposed locations. Under certain conditions of operation, i.e., pole-top mounting or platform mounting, an accident hazard is created if the operator must hold a flashlight in one hand to perform necessary tuning operation.

The crystal current and the modulation component of the wavemeter output may be monitored at the transmitter unit at jacks provided for the purpose. The video component of the off-the-air monitor must be amplified and transformed to a lower impedance for transmission through a coaxial cable to the control unit. A single stage of amplification, connected to a cathode follower for impedance matching, is used for this purpose.

Two stages of amplification are added to the single stage of the previous amplifier to obtain 32 db gain in the transmitter video amplifier. Feedback was used to stabilize the amplifier frequency and gain characteristics against tube variations and aging. The first two stages are connected in cascode to reduce plate current requirements and to simplify interstage coupling. Fig. 11 illustrates the circuit used. The video amplifier frequency response will remain flat within ±0.3 db from 15 cycles to 8 mc despite transconductance variations of 20 per cent in the video amplifier tubes. The bandpass of the amplifier between 6-db points is 3 cycles to 9.5 mc at 32 dbv gain. Fig. 12 illustrates typical gain-frequency runs obtained. Phase shift below 15 kc, as measured by Lissajous figures, is 0 at 20 cycles and rises to 11 degrees at 10 cycles and 23 degrees at 5 cycles. The output impedance of the amplifier is reduced to 900 ohms by the feedback, thereby reducing the effect of variations in klystron input and wiring capacities on the frequency response.

The d-c restorer is a *6AL5* miniature diode. Germanium diodes are unsatisfactory for this application because the back resistance of the crystals vary with changing operating temperature. Silicon junction diodes may make excellent d-c

Fig. 8 (left). Receiver grounding ring and finger stock

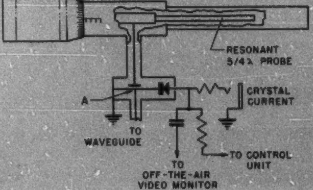


Fig. 10 (right). Coaxial wavemeter and crystal matching section

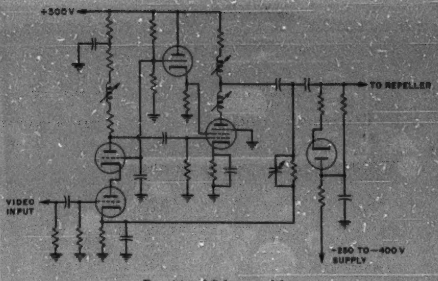


Fig. 11. Video amplifier

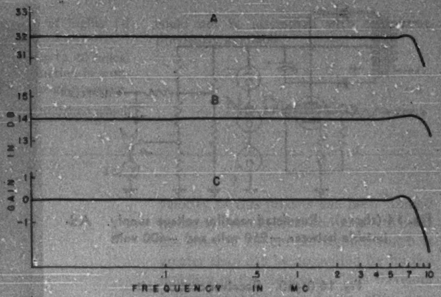


Fig. 12 Video gain-frequency

Curve A—Transmitter video amplifier
Curve B—Receiver video amplifier
Curve C—Over-all response

restorers when they are commercially available. D-c restorers are not used on systems transmitting *NTSC* color signals because the rise time of the leading edge of the sync pulse deteriorates if d-c restorers are employed with the equalizers used in this service application.

The transmitter unit interlock may be by-passed for maintenance purposes by rotating a butterfly device that holds down the interlock switch and simultaneously covers a hole through which a mounting stud must pass when the transmitter chassis is replaced in its case; see Fig. 7. This arrangement prevents positioning of the transmitter unit for service until the interlock action has been reactivated.

Two independent operator communication facilities are provided between the transmitter unit and the control unit. One facility consists of a pair of wires terminated in connecting posts. This facility is used with magneto sets and may be connected to an order wire at the control unit location. So arranged, an operator at the transmitter unit may talk to the operator at the receiving end of this microwave link during antenna alignment operations. The second facility is intended for use between the transmitter unit and control unit, and consists of an internal power supply and a pair of wires terminated at each end in standard double phone jacks. The power is obtained from the 6.3-volt filament supply, through a suitable rectifier and filter system.

TRANSMITTER CONTROL UNIT

The requirements that obtain with 1 watt output operation require the fabrication of a complete unit. Fig. 13 shows the new unit. The power requirements for the modified transmitter are: 750 volts ±1.0 volt at 0.08 ampere, 310 volts ±10 volts at 0.06 ampere, and a negative supply variable between 250 and 400 volts with a permissible variation of ±0.3 volt at any setting of the repeller voltage control. The limits are based on the requirement that the allowable power supply variations will not cause a frequency variation in excess of ±0.1 mc in the transmitting klystron.

A novel combination of circuits is used to attain the precise regulation necessary to maintain the desired frequency stability in the transmitting klystron. Fig. 14 is a schematic diagram of this voltage regulator circuit. Gas tube voltage regulators are used to derive the repeller voltage. It is characteristic of these tubes that the output voltage varies incrementally as the a-c supply changes. To circumvent this characteristic, a pentode with fixed control grid and screen grid voltages is placed in series with the voltage regulator tubes as a regulator for the current through the voltage regulator tubes. In effect the pentode acts as a current limiting resistance. As the a-c line voltage varies, the resistance of the pentode operates to maintain the current constant. The resultant regulation is on the order of 0.0025-per-cent repeller voltage change for each 1-per-cent change in line voltage, while 0.0054 per cent would suffice to meet the requirements just outlined.

The 750-volt supply for the accelerator anode of the klystron is regulated by a series tube regulating circuit to which has been added a secondary control feature that samples the unregulated voltage and adjusts the resistance of the series regulating tube to cancel out the effect of power line variations. The effectiveness of this circuit is such that variations in line voltage from 85 to 135 volts will result in less than a 0.4-volt change in the 750-volt supply. The 310-volt supply for the video amplifiers is derived from the 750-volt supply through a series resistance and is stabilized with voltage regulator tubes.

A display of the video signal obtained

Fig. 13. Transmitter control unit

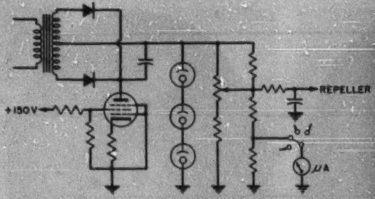


Fig. 14 (above). Regulated repeller voltage supply. Adjustable between −250 volts and −400 volts

Fig. 16 (right). Receiver control unit

from the r-f output of the transmitter is transmitted through coaxial cable to the control unit and is available for monitoring on the front panel as are the monitors for the video input and output cables. Complete metering facilities are provided for all regulated voltages, klystron current, power line voltage, and wavemeter crystal current

Input voltage variation in remote locations necessitated an adjustable, tapped input winding on the power transformer. To facilitate field adjustment, the taps are terminated in a switch located on the front panel. In practice, the tap that causes the line voltage meter to read correctly is selected.

An interlock has been provided on the rear cover panel of the power supply. This interlock may be bridged for maintenance purposes but it is so constructed that when the cover is replaced the interlock is automatically reactivated.

RECEIVER UNIT

The receiver unit is illustrated in Fig. 15. The *X26* klystron beating oscillator is coupled through a 20-db directional coupler to the main section of the wave guide. An adjustable probe in the directional coupler samples the beating oscillator frequency and is connected to a variable wavemeter as described in the section titled "Transmitter Unit." The beating oscillator excites the mixer crystal located in the end of the main section of the guide. Beating oscillator radiation is reduced to a minimum by this arrangement.

Fig. 15. Receiver unit

The klystron mechanical tuning drive is similar to the tuning drive in the transmitting unit in that approximately one turn is required to shift one channel and that the wavemeter and tuning drive operate on the frequency in the same manner when turned in the same direction.

The wavemeter in the receiver unit is identical to the transmitter wavemeter and may be exchanged with it if necessary. The wavemeter crystal current may be measured at both the receiver unit and the control unit for maximum operational flexibility.

I-f band-pass characteristics are critical in multiple link operations; therefore additional effort is expended to maintain a flat 20-mc bandpass. The loss of gain brought about by widening the i-f bandwidth was made up by increasing the screen voltage of the *6AK5* amplifier tubes. Noise levels of ten preamplifiers measured between 75 and 87 db below 1 milliwatt. Interference suppression measures have been included in the design.

The interlock and communications facilities provided on the receiver unit are identical to those on the transmitter unit.

RECEIVER CONTROL UNIT

Fig. 16 shows the receiver control unit. The new video amplifier is of feedback design to stabilize the frequency response and gain against tube variations and aging. The output impedance of the video amplifier was reduced to 75 ohms, thereby reducing reflections on the transmission line. A gain control that will not affect the frequency response has been added to the video amplifier. An input to the transmitter of 0.5 volt results in a 1.5-volt signal at the receiver when the receiver and transmitter gain controls are on the top step.

The bandpass of the main i-f amplifier must meet the requirements outlined for the i-f preamplifier in the receiver unit. Interference suppression features have been included in the design.

Transient interference introduced into the video output of the unmodifier system when the meter switch was operated was eliminated with circuit revisions which make possible in-service switching of the panel meter. Arrangements were made to monitor the wavemeter crystal current at the control unit. If the wavemeter is left "on frequency" then the receiver may be tuned to the correct frequency from the control unit by adjusting the repeller voltage and observing the dip in reading of the wavemeter current when the correct frequency is reached. This operating feature is useful when the receiver unit is located at a point remote from the control unit.

One-man operation of the receiver location has been made possible by providing controls for adjustment of the beating oscillator frequency at the receiver unit as well as at the receiver control unit.

Conclusions

The attempt to approach an ideal *SHF* relay system through the modification of an *RCA TTR/TRR* system has lead to the following conclusions.

1. It is possible to retain the desirable features of a *TTR/TRR* relay system when it is modified to meet the special needs of a communications company.

2. Significant savings in maintenance costs, operation costs, and plant investment are to be derived from the use of higher r-f power output.

3. Short-range operations (up to 40 miles) are simplified through the reduction in number of relay links required to give service.

4. Long facilities may be established which could not be attempted with equipment previously available.

◆

No Discussion

The Use of Sphere Gaps at Radio Frequencies

CHARLES B. OLER
MEMBER AIEE

FOR many years the sphere gap has been used for measuring high voltage at power frequencies[1,2] and it is also used for impulse testing. However, at present the sphere gap is not an accepted standard for the measurement of radio-frequency voltages. This is largely attributable to the lack of complete data on the spark-over values of sphere gaps in air, at radio frequencies. Some results have been published[3–6] but they are not altogether consistent and the spectrum of frequencies covered is far from complete.

This paper presents the results of studies on the spark-over voltage for 2-centimeter (cm) sphere gaps at frequencies from 300 kc to 12.5 megacycles (mc), with voltages ranging up to a 20-kv peak. Some data on 6.25-cm spheres, in the lower part of the frequency range, are also included. The results are shown in the curves in Figs. 1 to 5.

It is concluded that the sphere gap can serve as an accurate and reliable measuring device for voltages in the range of frequency covered. For best results the spheres should be irridiated by ultraviolet light or some other source of ionization should be provided. It is also found that the spark-over voltages show a transition from values which are essentially the same as those at power frequencies to values which are about 16 per cent lower. This change is attributed to an accumulation of ions in the gap at higher frequencies and larger gap values and it is estimated from the data that the average ion mobility in air at normal pressure is approximately 8 cm per second per volt per cm.

The Use of Sphere Gaps at Power Frequencies

The measurement of high voltage at power frequencies can readily be made with good accuracy by using sphere gaps. Tabulated values of spark-over voltages in air for spheres of various diameters and spacings are given in reference 1. A simple correction for air density must be made, because the dielectric strength of air is roughly proportional to its density. It should be noted that the sphere gap is a peak-reading or crest voltmeter. Although the gap spacing is critical in determining the spark-over voltage, the size of the spheres is not, so long as the gap is small compared to the diameter of the spheres. Under these conditions the electric field in the gap, along the axis of the spheres, is fairly uniform. When used for measuring voltage, sphere gaps have the disadvantage that the spark-over imposes practically a short circuit on the voltage source. Consequently it is customary to connect a fairly high resistance in series with the sphere gap.

Theory of Spark-Over

The generally accepted theory of the spark-over of a gap[7] postulates that electrons existing in the gap are attracted toward the anode and, if the electric field intensity is great enough, they acquire sufficient energy during the traverse of a mean free path to ionize a molecule on collision. This produces more electrons and an "avalanche" is formed which sweeps toward the anode and leaves in its trail a highly ionized column of air. The velocity of this process has been studied by many investigators and found to be of the order of 10^8 cm per second, which is greater than can be accounted for by the mobility of the electrons. It is therefore believed that the ionized core of the avalanche emits light photons, some of which cause ionization in advance of the avalanche. Spark-over occurs when an ionized path completely bridges the gap.

Effects of Irradiation

The initiation of the spark-over mechanism requires the existence of some electrons in the gap. Hence there is some uncertainty in the value of voltage at which a gap will break down, unless the voltage is maintained for a reasonable time. In other words, the process involves an unpredictable time lag. To reduce this uncertainty it is sometimes the practice to irradiate the sphere gap, using ultraviolet light, X rays, or emissions from radioactive materials. A generous supply of electrons in the gap is thus assured, the spark-over time lag is reduced, the the observed values of breakdown voltage are

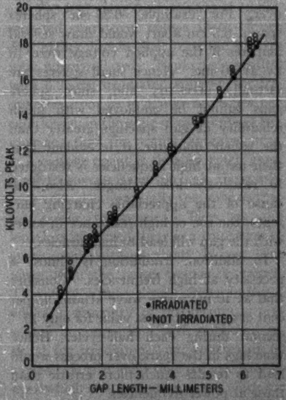


Fig. 1. Sphere gap spark-over voltages at 300 kc. Data include 2- and 6.25-cm spheres, with no perceptible difference in results. Sphere gap vertical with the lower sphere grounded

---

Paper 54-294, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 22, 1954; made available for printing May 10, 1954.

Charles B. Oler is with the United States Naval Postgraduate School, Monterey, Calif.

The author gratefully acknowledges his indebtedness to Dr. C. F. Miller of The Johns Hopkins University for inspiration and help in this study.

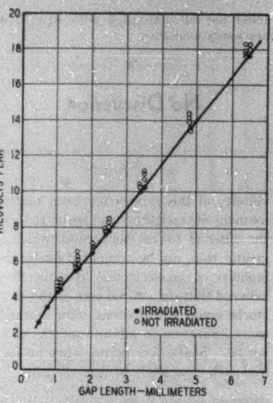


Fig. 2. Sphere gap spark-over voltages at 750 kc. Data include 2- and 6.25-cm spheres

more consistent. Irradiation generally lowers the spark-over voltage somewhat.[8] At power frequencies, it is usually considered necessary only for short gaps.

Use of Sphere Gaps at High Frequencies

Some modification of practice and theory is needed when the sphere gap is used at high frequencies. In the first place, spheres of large diameter are unsuitable because of the excessive charging current they would draw prior to spark-over. For example, 6.25-cm spheres spaced 0.5 cm apart would draw several amperes if the applied voltage were 10 kv at 10 mc. Hence small spheres are virtually necessary and, since sphere gaps cannot be employed with much reliability at gap spacings greater than the sphere diameter, it is evident that their use at high frequencies is restricted to relatively low voltages. Also, because of the appreciable charging currents, the use of high resistance in series with the gap will lead to inaccuracies.

In addition, irradiation is almost a necessity at high frequencies. Consider that at 10 mc a sine wave attains 95 per cent or more of its crest value for only 10^{-8} second during each half-cycle. Hence time lags in the spark-over process would lead to results much more erratic than those at power frequencies.

The avalanche theory of the spark-over is applicable so long as the electrons encounter many collisions during each cycle of applied voltage. However, if the frequency is so high or the gas pressure so low that the electrons experience many cycles of applied field between collisions, then this theory is not considered applicable.[7] In air at atmospheric pressure electrons will encounter many collisions during each cycle if the frequency is less than approximately 10^{10} cycles per second. Hence theoretical considerations of radio-frequency spark-over at atmospheric pressure can be based on the avalanche concept.

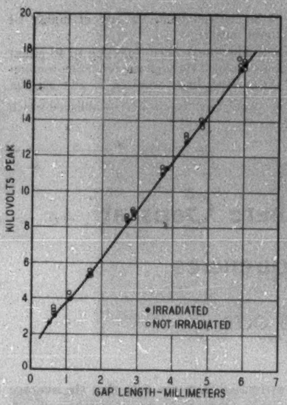


Fig. 3. Sphere gap spark-over voltages at 2,088 kc. Data include 2- and 6.25-cm spheres

Experimental Procedure

To test the performance of sphere gaps at radio frequencies it was necessary to devise a suitable method of accurately measuring the voltages involved. For this purpose a vacuum-tube voltmeter, with associated capacitance divider, was constructed and calibrated. Its error is estimated to be less than 2 per cent. This type of voltmeter is particularly suitable because it is a peak-reading instrument. Calibration of the voltmeter was accomplished by simultaneous readings of the voltmeter and of the current drawn by a known capacitor, when these were both connected to the high-voltage source. A thermocouple and associated microammeter previously calibrated on direct current were used to measure the capacitor current. The voltmeter was calibrated up to a 20-kv peak, using frequencies ranging from 300 kc to 5 mc. No variation with frequency was found.

A schematic diagram of the circuit arrangement used for the sphere gap tests is shown in Fig. 6, which also shows the basic circuit of the voltmeter. Capacitances C_1 and C_2 comprise a voltage divider. C_1 is an air capacitance formed by two disks approximately 4 inches and 1 inch in diameter respectively and mounted $1^1/_3$ inches apart. The smaller disk is surrounded by a guard ring. The value of C_1 is calculated to be 0.093 micromicrofarads (mmf). C_2 is a 25-mmf mica capacitor, plus the additional capacitance to ground of the smaller disk and the necessary wiring. The value of C_2 was found by measurement to be 33 mmf. The ratio of the voltage divider is therefore approximately 350 to 1. A type-*955* vacuum tube is used in the voltmeter with grid and plate tied together to form a diode. The resistance *R* is 3 megohms.

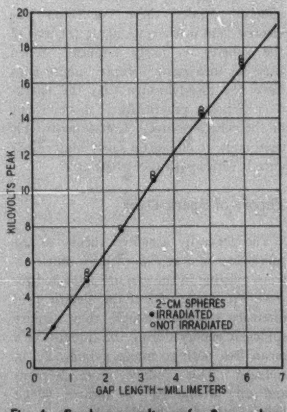


Fig. 4. Spark-over voltages for 2-cm sphere at 4,873 kc

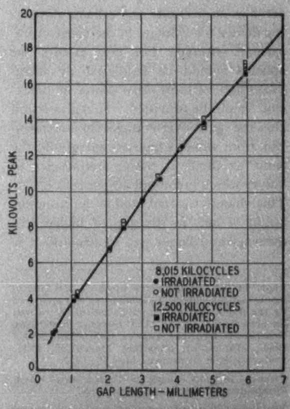


Fig. 5. Spark-over voltages for 2-cm spheres at 8,015 and 12,500 kc

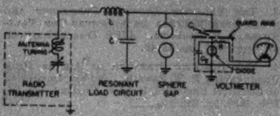


Fig. 6. Schematic circuit of the testing setup

The time constant RC_2 is approximately 100 microseconds, permitting use of the instrument at frequencies as low as 300 kc without appreciable error due to decay of the voltage across C_2 between cyclic peaks. Also, the susceptance of C_2 at 300 kc is approximately 200 times the conductance of R_2, so the ratio of the capacitance divider is negligibly altered by the shunting effect of R_2. A 0–10 microammeter having 1/2 per cent or less error is used to measure the diode current. The over-all effective capacitance of the voltmeter is approximately 5 mmf, attributable mostly to the capacitance between the guard ring and the 4-inch disk mounted $1^1/_2$ inches above it.

Two navy-type radio transmitters served as sources of test voltage. One was rated 1 kw from 300 kc to 2 mc, and the other was rated 500 watts, from 2 to 18 mc. The high voltages were obtained by using a series-resonant inductance-capacitance load circuit. This arrangement kept the harmonics at a minimum and wavemeter tests consistently showed negligible harmonic content in the output. The voltmeter and sphere gap were connected in parallel with the capacitance section of the inductance-capacitance load circuit, which further reduced the possibility of harmonic distortion in the test voltage.

The load inductance L was any one of the three air-core coils, which had inductances of 500, 100, and 10 microhenrys respectively. For the lowest frequency used the load capacitance C was a 500-mmf mica capacitor. At the highest frequency the parallel capacitance to ground of the voltmeter, the sphere gap, and the connecting leads was sufficient to provide resonance with the 10-microhenry coil. For intermediate frequencies various values of C were simply constructed by mounting a metallic plate of appropriate size several inches above a ground plane.

Two sphere gaps were used. The larger was a commercial type having 6.25-cm brass spheres. The smaller was specially constructed for this study, having 2-cm brass spheres supported in a polystyrene and plywood frame. In all tests the sphere gap was in a vertical position, with the lower sphere grounded.

A simple source of irradiation was chosen and found to be adequate for the purpose. It was an ordinary mercury-arc-type sun lamp which fits into a standard lamp socket. The glass bulb is a special kind which passes considerable ultraviolet light. The effect of this type of irradiation is to release photo electrons from the surface of the spheres.

To get the test data a selected frequency was set up in the transmitter and a suitable coil and capacitance were arranged to provide a series-resonant load. The capacitance included that of the sphere gap and the voltmeter. Within the transmitter were a variable loading coil and a variable coupling capacitor which permitted a smooth, vernier final adjustment for resonance. The sphere gap was then set at a definite gap spacing and the plate voltage on the output tubes of the transmitter was increased until a preliminary spark-over occurred. The voltage was then cut off and the load detuned slightly so that when plate voltage was restored to the output tubes the radio-frequency voltage across the sphere gas was about 90 per cent of spark-over value. This voltage was slowly increased by restoring the resonance, using gradual adjustment of the antenna loading coil or the coupling capacitor. Spark-over normally was then reached in approximately 10 to 20 seconds. The transmitter was cut off manually as soon as spark-over took place, to avoid burning or pitting the spheres. Five such readings were taken and averaged for each data point, without irradiation of the spheres. The ultraviolet light was then turned on and five more readings were taken. Finally, the gap setting was rechecked before being reset to a new spacing.

The atmospheric pressure was read from a mercury-column barometer. Temperature and humidity were recorded from a wet-and-dry bulb thermometer. The spheres were cleaned frequently with crocus cloth, carbon tetrachloride, and a soft cloth, although cleaning seemed to produce no noticeable effect on the results.

Results and Conclusions

The results are presented in the curves in Figs. 1 to 5. Each point shown is the average of five successive readings, corrected to standard conditions of pressure and temperature (760-millimeter mercury pressure and 25 degrees centigrade). The principal conclusion is that there is very little variation in the results over the range of frequencies covered, from 300 kc to 12.5 mc. With the spheres irradiated the results are very consistent and repeatable. If the spheres are not irradiated, the consistency is not as good, and the spark-over voltages are generally somewhat higher.

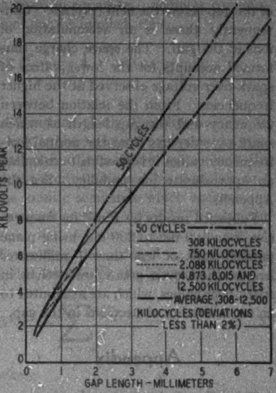


Fig. 7. Summary of spark-over tests on 2-cm irradiated sphere gap. The 50-cycle data are from previously published results[9]

The results are in general agreement with others previously published, which cover part of this frequency range. Ekstrand's results[6] with 2-cm irradiated spheres at 700 and 1,800 kc agree very closely indeed. Seward,[4] using 1.4-cm spheres at 100, 600, and 900 kc, and Alford and Pickles,[5] using 5-cm spheres at 13 mc, reported spark-over values 10 to 15 per cent higher, but in their tests the spheres were not irradiated. Reukema's results[3] with 6.25-cm irradiated spheres, for frequencies from 28 to 425 kc are a few per cent lower, in the zone of overlapping frequencies, than the results presented here.

A curious result is that for small gaps, if the frequency is not too high, the spark-over voltage is essentially the same as that for power frequencies. Then as gap length is increased there is a transition to spark-over voltages about 16 per cent lower than the power-frequency values. This effect is most clearly shown in Fig. 7 where a summary of the results is presented in comparison with published results[9] for irradiated 2-cm spheres at 50 cycles. It is evident that the gap lengths over which this transition takes place are roughly inversely proportional to the frequency. This effect is attributed to the theory that with higher frequency and longer gap length the ions formed during a half-cycle of voltage are not swept out of the gap before the field reverses. Consequently there is an accumulation of ions in the gap. The space charge thus formed accounts for the lower values of spark-over voltage observed at the higher frequencies. From the relation between frequency and the gap length at which spark-over departs from the normal low-frequency values it is possible to calculate an approximate ion mobility. See the Appendix.

Published results by Pim[10] for frequencies from 100 mc to 300 mc, using plane parallel gaps of less than 1 millimeter length, show even greater reductions in spark-over values, which are attributed to an accumulation of electrons in the gap.

Appendix

The drift velocity v of an ion in a gas, under the influence of an electric field X volts per cm, is usually taken to be $v=kX$ where k is the ion mobility expressed in cm per second per volt per cm. Assuming that ions are formed at the peak of an impressed sinusoidal electric field, their subsequent velocity would be $v=kE\cos 2\pi ft$ where E is the peak field strength and f is the frequency. During the next quarter-cycle, until the field diminishes to zero, the ions will move a distance

$$d=\int_0^{\frac{1}{4f}} v\,dt=\frac{kE}{2\pi f}\text{ cm}$$

hence

$$k=\frac{2\pi fd}{E}$$

Now, from Fig. 7, it can be concluded that at a frequency of 308 kc the ions are swept from the gap, for gap values of 1.25 millimeters or less. Taking E to be 30,000 volts per cm, which is approximately the spark-over value for air, gives

$$k=\frac{2\pi(308{,}000)(0.125)}{30{,}000}\doteq 8\text{ cm per second per volt per cm}$$

References

1. MEASUREMENT OF TEST VOLTAGE IN DIELECTRIC TESTS. *AIEE Standard No. 4*, Mar. 1943 (Revised Sept. 1953).

2. DIELECTRIC PHENOMENA IN HIGH VOLTAGE ENGINEERING (book), F. W. Peek, Jr. McGraw-Hill Book Co., Inc., New York, N. Y., 1929.

3. THE RELATION BETWEEN FREQUENCY AND SPARK-OVER VOLTAGE IN A SPHERE-GAP VOLTMETER, L. E. Reukema. *AIEE Transactions*, vol. 47, Sept. 1927, pp. 38–49.

4. THE ELECTRIC STRENGTH OF AIR AT HIGH FREQUENCIES, E. W. Seward. *Journal*, Institution of Electrical Engineers, London, England, vol. 84, Feb. 1939, pp, 288–94.

5. RADIO-FREQUENCY HIGH-VOLTAGE PHENOMENA, Andrew Alford, Sidney Pickles. *AIEE Transactions (Electrical Engineering)*, vol. 59, Mar. 1940, pp. 129–37.

6. RADIO-FREQUENCY SPARK-OVER IN AIR, P. A. Ekstrand. *Proceedings*, Institute of Radio Engineers, New York, N.Y., vol. 28, June 1940, pp. 262–66.

7. FUNDAMENTAL PROCESSES OF ELECTRICAL DISCHARGE IN GASES (book), L. B. Loeb. John Wiley and Sons, Inc., New York, N. Y., 1939.

8. THE EFFECTS OF IRRADIATION, HUMIDITY, AND SPHERE MATERIAL ON THE SPARK-OVER VOLTAGE OF THE TWO-CENTIMETER SPHERE GAP, A. B. Lewis. *Journal of Applied Physics*, New York, N. Y., vol. 10, Aug. 1939, pp. 573–77.

9. THE CALIBRATION OF 2-CM. DIAMETER SPHERE GAPS, R. Cooper, D. E. M. Garfitt, J. M. Meek. *Journal*, Institution of Electrical Engineers, London, England, vol. 95, pt. II, June 1948, pp. 309–11.

10. THE BREAKDOWN STRENGTH OF AIR AT ULTRA-HIGH FREQUENCIES, J. A. Pim. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 96, pt. III, 1949, pp. 117–29.

◆

No Discussion

Analysis and Synthesis of Sampled-Data Control Systems

ELIAHU I. JURY
ASSOCIATE MEMBER AIEE

Synopsis: The theory of sampled-data systems using the method of the z-transform is extended and clarified. In particular, the equivalence between the z-transform in its closed form and the infinite summation used by some investigators is shown. Important characteristics of the pulsed transfer function, and initial and final value theorems are developed for the z-domain. An extensive table is given of z-transform pairs covering the most important and most commonly encountered system functions and input functions. The technique for stabilizing and shaping the pulsed transfer locus is demonstrated. In particular, the application of linear compensating networks in the continuous part of the system is investigated. Design criteria are obtained which relate the transient response of sampled-data systems and the frequency response.

CONTROL systems having signals applied at equally spaced intervals of time at one or more points of the system and which receive no information between two consecutive signals are known as sampled-data control systems. Such systems are introduced when information to the system is received as intermittent data, and sometimes their introduction is required to gain sensitivity, at the expense, however, of bandwidth.

The analyses of sampled-data control systems have attracted the attention of many investigators in the field, both in this country and in Europe. Basically, the analyses applied by most writers involve the use of the Laplace transform. The objectives of such analyses are twofold. One concerns itself with the output at sampling instants, where the z-transform method is best applicable, while the other seeks to obtain the output as a continuous function of time using ordinary Laplace transforms.

Synthesis and design of sampled-data control systems has not been thoroughly investigated where quantitative design information can be achieved. Linvill's[1] investigation was the first attempt in the field to obtain design information. However, the results obtained are approximate, and unless checked by the z-transform method, the validity of approximation is in doubt.

Based on the z-transform method of analyses,[2–5] this paper introduces certain techniques that make possible the application of this method to system synthesis and design. The material introduced in the analysis is mainly concerned with the clarification and further extension wherever possible of the z-transform method. In particular, it is shown that the equivalence of the two different forms (the infinite summation and the closed form) of sampled-data system transfer function is obtainable by contour integration in the right as well as the left half of the complex frequency plane p. Furthermore, the important characteristics of this transfer function are pointed out. Con-

---

Paper 54-298, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted December 22, 1953; made available for printing May 13, 1954.

ELIAHU I. JURY is with the University of California, Berkeley, Calif.

Appreciation is gratefully expressed to Prof. John R. Ragazzini for his help and valuable suggestions and for the time he devoted to guiding the progress of this paper. Thanks are also expressed to Profs. John B. Russell and L. A. Zadeh. This research was supported in part by the United States Air Force under Contract No. *AF 18(600)677*, monitored by the Office of Scientific Research Air Research and Development Command.

This material is part of the thesis submitted for the degree of Doctor of Engineering Science at Columbia University, New York, N. Y.

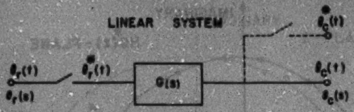


Fig. 1. Sampled-data system

ditions for initial and final value theorems in the z-domain are developed. The extensive list introduced in Table I is useful in readily obtaining the sampled-output response of such systems when subjected to various forms of inputs.

Applying the circle locus method of mapping the system transfer function in the z-plane, the problem of shaping is much simplified. Furthermore, confining the shaping of transfer loci to insertion of linear networks (although digital computers and pulsed networks can be used), the important effects of such networks are represented.

In correlating the graphical-mathematical relationship between the shape of the over-all transfer loci to the transient response of sampled-data control systems, a relationship between the maximum of the frequency response and the maximum of the sampled response of second-order systems is developed.

Mathematical Background

SAMPLED-DATA SYSTEMS TRANSFER FUNCTIONS

Just as continuous systems are characterized by a transfer function which relates the transform of the continuous output to the input, so can a sampled-data system shown in Fig. 1 be described by a transfer function which relates the transform of the sampled output to the input. This sampled-data system transfer function is denoted as $G^*(\epsilon^{sT})$ and defined as the Laplace transform of the product of the impulsive response of the linear system and chain of unit impulses repeated at equal interval T, thus

$$G^*(\epsilon^{sT})=\mathfrak{L}[g(t)\delta_T(t)] \tag{1}$$

where $g(t)$ represents the impulsive response of the linear system, whose transform is $G(s)$, and $\delta_T(t)$ represents a sequence of unit impulses repeated at equal intervals T, and extended from minus infinity to plus infinity. Since for relaxed systems

$$g(t)=0 \text{ for } t<0 \tag{2}$$

$\delta_T(t)$ in equation 1 need be considered only for positive values of time. Hence its transform can be expressed as

$$\mathfrak{L}[\delta_T(t)]=1+\epsilon^{-Ts}+\epsilon^{-2Ts}+\ldots=\frac{1}{1-\epsilon^{-Ts}} \tag{3}$$

Evaluating $G^*(\epsilon^{sT})$ in terms of $G(s)$ as given in equation 1, complex convolution[6] can be applied, thus

$$G^*(\epsilon^{sT})=\frac{1}{2\Pi j}\int_{c-j\infty}^{c+j\infty}G(p)\frac{1}{1-\epsilon^{-T(s-p)}}dp \tag{4}$$

Integration along the line $c-j\infty$ to $c+j\infty$ in the p-plane is equivalent to integration in the negative sense along the closed contour formed by that line and the infinite semicircle which encloses the singularities of $1/(1-\epsilon^{-T(s-p)})$ in the right half p-plane as shown in Fig. 2. Evaluating this integral in this contour yields for $G^*(\epsilon^{sT})$ the following

$$G^*(\epsilon^{sT})=\sum_{n=-\infty}^{n=\infty}\frac{G(s+jn\omega_r)}{T} \tag{5}$$

where

$=$ integer

$$\omega_r=\frac{2\Pi}{T} \tag{6}$$

Equation 4 is useful in obtaining an approximate relation for $G^*(\epsilon^{sT})$ when certain conditions, such as when T is small or when $G(s)$ is of low pass form,[1] exist.

An alternate form of $G^*(\epsilon^{sT})$ can be derived by differently evaluating the integration in equation 3. For the integration along the line $c-j\infty$ to $c+j\infty$ is

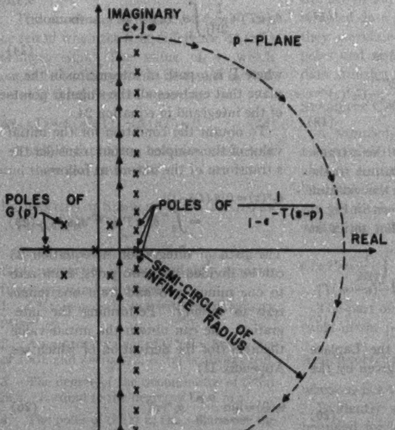

IMAGINARY
c+j∞
p-PLANE
POLES OF G(p)
POLES OF $\frac{1}{1-\epsilon^{-T(s-p)}}$
REAL
SEMI-CIRCLE OF INFINITE RADIUS
c-j∞

Fig. 2. Path of integration in the right half-plane of complex frequency p

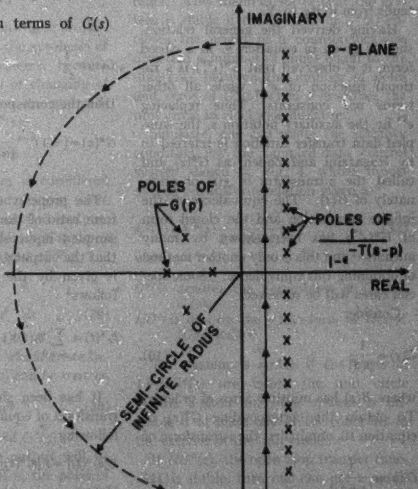

IMAGINARY
p-PLANE
POLES OF G(p)
POLES OF $\frac{1}{1-\epsilon^{-T(s-p)}}$
REAL
SEMI-CIRCLE OF INFINITE RADIUS

Fig. 3. Path of integration in the left half-plane of complex frequency p

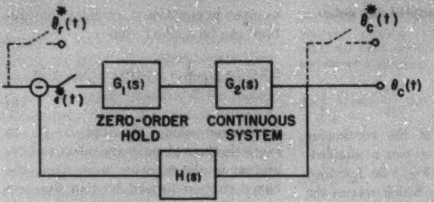


Fig. 4. Typical sampled-data control system

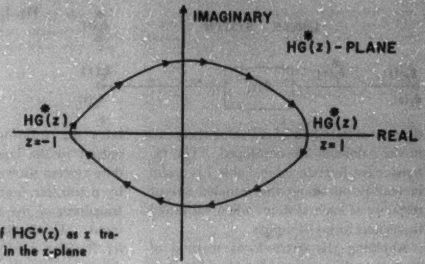


Fig. 5 (right). Contour of HG*(z) as z traverses the unit circle in the z-plane

also equivalent to integration in a positive sense along the closed contour formed by that line and the infinite semicircle in the left half-plane of complex frequency p, as shown in Fig. 3. Assuming

$$G(s)=\frac{A(s)}{B(s)} \tag{7}$$

where $A(s)$ and $B(s)$ are polynomials in (s), and $B(s)$ includes only simple poles and is of higher order than $A(s)$, $G^*(\epsilon^{sT})$ can be easily shown to be

$$G^*(\epsilon^{sT})=\sum_{n=1}^{N}\frac{A(s_n)}{B'(s_n)}\frac{1}{1-\epsilon^{-T(s-s_n)}} \tag{8}$$

where

$$B'(s_n)=\left.\frac{dB}{ds}\right|_{s=s_n} \tag{9}$$

and s_n is a typical pole of $G(s)$ and extends from unity to N.

Having derived the general relationship as given in equation 8, in a closed form, it is observed that $G^*(\epsilon^{sT})$ is a rational function of ϵ^{Ts}, while all other terms are constants. Thus replacing ϵ^{Ts} by the auxiliary notation z, the sampled-data transfer function is referred to by Ragazzini and Zadeh[6] as $G^*(z)$ and called the z-transform of $g(t)$ or alternately of $G(s)$. The equivalence of the infinite summation and the closed form of $G^*(\epsilon^{sT})$ has been shown by many authors[2,6] and this is only another method of obtaining this equivalence. Some special cases will be reviewed.

Consider

$$G(s)=\frac{1}{B(s)} \tag{10}$$

where $B(s)$ has multiple zeros of order n. To obtain the corresponding $G^*(z)$ of equation 10, obtain first the z-transform of

$$G_1(s)=\frac{1}{(s+a)^2} \tag{11}$$

Equation 11 can be written as

$$G_1(s)=-\frac{\partial}{\partial a}\left(\frac{1}{s+a}\right) \tag{12}$$

When the z-transformation of both sides of equation 12 is effected, then

$$G_1^*(z)=-\frac{\partial}{\partial a}G_0^*(z) \tag{13}$$

where $G_0^*(z)$ is the z-transform of $1/(s+a)$ which can readily be obtained with the use of relation 8. Similarly, if

$$G(s)=\frac{1}{(s+a)^n} \tag{14}$$

then

$$G(s)=(-1)^{n-1}\frac{1}{(n-1)!}\frac{\partial^{n-1}}{\partial a^{n-1}}\left(\frac{1}{s+a}\right) \tag{15}$$

and the corresponding z-transform of equation 14 is

$$G^*(z)=(-1)^{n-1}\frac{1}{(n-1)!}\frac{\partial^{n-1}}{\partial a^{n-1}}G_0^*(z) \tag{16}$$

If

$$G(s)=\frac{1}{s^n} \tag{17}$$

then the corresponding $G^*(z)$ is

$$G^*(z)=(-1)^{n-1}\frac{1}{(n-1)!}\lim_{a\to 0}\frac{\partial^{n-1}}{\partial a^{n-1}}G_0^*(z) \tag{18}$$

The property of $G^*(z)$ as the z-transform ratio of the sampled output to the sampled input stems from the relation that the output $\theta_c^*(t)$, as shown in Fig. 1 is given by the superposition series as follows[3]

$$\theta_c^*(t)=\sum_{m=0}^{\infty}\theta_r(t)\delta(t-mT)g(t)\delta[t-(n-m)T] \tag{19}$$

It has been shown[7] that the Laplace transform of equation 19 is given by the following

$$\mathfrak{L}[\theta_c^*(t)]=\mathfrak{L}[\theta_r^*(t)]\mathfrak{L}[g^*(t)] \tag{20}$$

Hence

$$\mathfrak{L}[g^*(t)]=\frac{\mathfrak{L}[\theta_c^*(t)]}{\mathfrak{L}[\theta_r^*(t)]} \tag{21}$$

Since[5]

$$\begin{aligned}\theta_c^*(t)&=\theta_c(t)\delta_T(t)\\ \theta_r^*(t)&=\theta_r(t)\delta_T(t)\\ g^*(t)&=g(t)\delta_T(t)\end{aligned} \tag{22}$$

then equation 21 can be written

$$G^*(z)=\frac{\theta_c^*(z)}{\theta_r^*(z)} \tag{23}$$

Therefore the sampled-data transfer function $G^*(z)$ relates the z-transform of the output to that of the input.

REAL INVERSION OF Z-TRANSFORM, INITIAL AND FINAL VALUE THEOREMS

The z-transform approach yields the output only at sampling instants, through a mathematical process known as real inversion. The real output at sampling instants can be obtained as the coefficient of the complex Fourier series expansion of $\theta_c^*(z)$, expressed by the following real inversion equation[8]

$$\theta_c(nT)=\frac{1}{2\Pi j}\int_{\Gamma}\theta_c^*(z)z^{n-1}dz \text{ (see footnote)} \tag{24}$$

where Γ is a path of integration in the z-plane that encloses all the singular points of the integrand in equation 24.

To obtain the condition for the initial value of the sampled output, consider the z-transform of the output as follows

$$\begin{aligned}\theta_c^*(z)&=\mathfrak{L}[\theta_c(t)\delta_T(t)]\\ &=\int_0^{\infty}\theta_c(nT)\epsilon^{-nTs}d(nT)\end{aligned} \tag{25}$$

The path of integration in equation 25 can be divided into two parts, from zero to one minus zero and from one minus zero to infinity. Performing the integration, one can obtain the initial value theorem (for the derivation of which see Appendix II)

$$\theta_c(0)=\lim_{z\to\infty}\frac{z-1}{z}\theta_c^*(z) \tag{26}$$

In this paper the real output at sampling instants is denoted as $\theta_c(nT)$ and is related to the output representation $\theta_c^*(t)$ in the following relation[5]

$$\theta_c^*(t)=\sum_{n=-\infty}^{n=\infty}\theta_c(nT)\delta(t-nT)$$

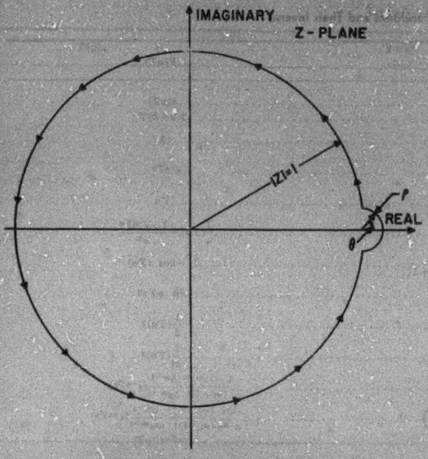


Fig. 6 (above). Closed path in the z-plane

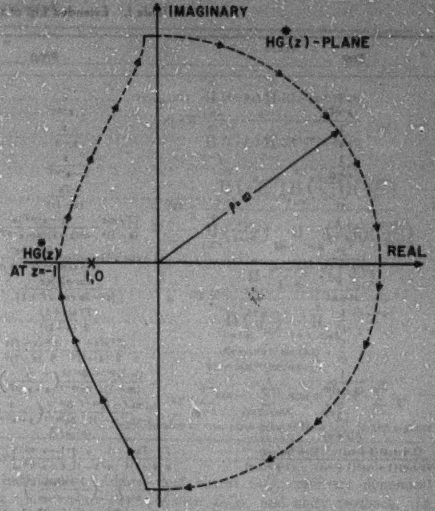


Fig. 7 (right). Contour of HG*(z) locus as z traverses a path shown in Fig. 6

where $\theta_c(nT)=0$ for $n<0$. In a similar manner one may obtain the final value theorem

$$\lim_{n\to\infty}[\theta_c(nT)]=\lim_{z\to 1}\left[\frac{z-1}{z}\,\theta_c^*(z)\right] \tag{27}$$

if all poles of $\theta_c^*(z)$ are inside the unit circle

The maximum ordinate of $\theta_c(nT)$ can be found when the first difference $\Delta\theta_c(nT)$ changes sign. The value of n, which ought to be integer, can be found by first solving the difference equation

$$\Delta\theta_c(nT)=\theta_c(n+1)T-\theta_c(nT)=0 \tag{28}$$

Table I lists the z-transform relations and their real inversion.

Properties of Sampled-Data Transfer Function

From equation 8 the following important properties of sampled-data transfer functions can be deduced:

1. $G^*(z)$ is periodic in s with the imaginary period $j\omega_r$.

2. Asymptotic values of $G^*(z)$—For values of $s=0$, and $s=jn\omega_r/2$, $G^*(s)$ is always real.

3. The degree of the denominator of $G^*(z)$ in z^{-1} is equal to the degree of $G(s)$ in s.

4. The poles of $G^*(z)$ in the s-plane are the same as those of $G(s)$.

5. Changing the values of the poles of $G(s)$ changes the coefficients $A(s_n)/B'(s_n)$ as well as the terms $1/(1-\epsilon^{-T(s-s_n)})$ in $G^*(z)$.

6. Insertion of zeros in $G(s)$ changes the coefficients $A(s_n)/B'(s_n)$ alone.

7. Insertion of poles in $G(s)$ increases the number of terms of $G^*(z)$ in the summation of equation 8; that is, N becomes larger and $G^*(z)$ will have more terms of the form

$$\frac{A(s_n)}{B'(s_n)[1-\epsilon^{-T(s-s_n)}]}$$

It will be shown later that these properties are of importance in the synthesis of sampled-data control systems, because they correlate the effects of changing of poles and zeros of $G(s)$ on the sampled-data transfer function $G^*(z)$.

Stability Considerations

A sampled-data system is considered to be stable if to a bounded input there always corresponds a bounded output. The stability conditions for such systems can be obtained if the output of Fig. 1 is written as follows

$$\theta_c(nT)=\frac{1}{2\Pi j}\int_{\Gamma} G^*(z)\theta_r^*(z)z^{n-1}dz \tag{29}$$

The part of the output which arises because of the poles of $G^*(z)$ can be written

$$[\text{residue of } G^*(z)]\times\theta_r^*(z_k)z_k^{n-1} \tag{30}$$

where z_k is a typical pole of $G^*(z)$.

Evidently so that the output will be bounded for any value of n, the poles of $G^*(z)$ must satisfy

$$|z_k|<1 \tag{31}$$

Therefore, for the sampled-data system to be stable all the poles of $G^*(z)$ should lie inside the unit circle.

For the case that the zeros of $G^*(z)$ lie inside the unit circle in the z-plane, one can apply a form of Nyquist criterion which is based on one of Cauchy's theorems. This theorem states that the system is stable if the contour of $G^*(z)$ encloses the origin in a counterclockwise direction a number of times equal to the difference between the number of poles and zeros of $G^*(z)$ when z describes the unit circle in clockwise direction.

Similarly, one can relate the stability condition for the sampled-data control system shown in Fig. 4, whose input-output z-transform relation given as[5]

$$\theta_c^*(z)=\theta_r^*(z)\frac{G^*(z)}{1+HG^*(z)} \tag{32}$$

where

$G^*(z)$ denotes the z-transform of $G_1(s)\times G_2(s)$

$HG^*(z)$ denotes the z-transform of $H(s)\times[G_1(s)\times G_2(s)]$

The system is stable if the roots of $1+HG^*(z)$ are inside the unit circle in the z-plane and the system is unstable if at least one of the roots lies in the exterior of the unit circle.

If $HG^*(z)$, the open loop transfer function, is stable, then one can use a modified form of a Nyquist criterion which states that if the transfer locus $HG^*(z)$ does not enclose the point $(-1,0)$ in the $HG^*(z)$-plane, then the system is stable, and vice versa.[3,11]

Table I. Extended List of z-Transforms and Their Inverse

F(s)	F*(z)	F(nT)
1	1	$\delta(nT)$
ϵ^{-kTs}	z^{-k}	$\delta(n-k)T$
$\frac{1}{s}$	$\frac{z}{z-1}$	1
$\frac{1}{s+a}$	$\frac{z}{z-\epsilon^{-aT}}$	ϵ^{-aTn}
$\frac{1}{s^2}$	$\frac{Tz}{(z-1)^2}$	Tn
$\frac{1}{s^2(s+a)}$	$\frac{(T/a)z}{(z-1)^2}-\frac{(1-\epsilon^{-aT})z}{a^2(z-1)(z-\epsilon^{-aT})}$	$\frac{1}{a}(Tn)-\frac{1-\epsilon^{-aTn}}{a^2}$
$\frac{a^2}{s(s^2+a^2)}$	$(1-\cos aT)\frac{(1+z)z}{(z-1)[z^2-2z\cos aT+1]}$	$(1-\cos aTn)$
$\frac{a}{s^2+a^2}$	$\frac{\sin aT\times z}{[z^2-2z\cos aT+1]}$	$\sin(cTn)$
$\frac{1}{s^3}$	$\frac{T^2}{2}\frac{z(z+1)}{(z-1)^3}$	$\frac{1}{2}(Tn)^2$
$\frac{1}{s^4}$	$\frac{T^3}{2}\frac{z(z+1)}{(z-1)^4}+\frac{T^3}{6}\frac{(z+2)z}{(z-1)^3}$	$\frac{1}{3!}(Tn)^3$
$\frac{1}{(s+a)^m}$	$\frac{(-1)^{m-1}}{(m-1)!}\frac{\partial^{m-1}}{\partial a^{m-1}}\left(\frac{z}{z-\epsilon^{-aT}}\right)$	$\frac{(-1)^{m-1}}{(m-1)!}\frac{\partial^{m-1}}{\partial a^{m-1}}(\epsilon^{-aTn})$
$\frac{1}{s^m}$	$\lim_{a\to 0}\frac{(-1)^{m-1}}{m-1!}\frac{\partial^{m-1}}{\partial a^{m-1}}\left(\frac{z}{z-\epsilon^{-aT}}\right)$	$\lim_{a\to 0}\frac{(-1)^{m-1}}{(m-1)!}\frac{\partial^{m-1}}{\partial a^{m-1}}(\epsilon^{-aTn})$
$F(s+a)$	$F(\epsilon^{aT}z)$	$\epsilon^{-aTn}F(nT)$
$\frac{(1+\tau_1 s)(1+\tau_3 s)\ldots(1+\tau_{2m-1}s)}{s^2(s+a)(1+\tau_2 s)(1+\tau_4 s)\ldots(1+\tau_{2m}s)}$	$\frac{1}{a}\frac{Tz}{(z-1)^2}-\frac{1}{a^2}\frac{z}{z-1}\left[\frac{1-\epsilon^{-aT}}{z-\epsilon^{-aT}}\right]\times$ $\left(\frac{1-a\tau_1}{1-a\tau_2}\right)\ldots\left(\frac{1-a\tau_{2m-1}}{1-a\tau_{2m}}\right)$ $-\frac{\tau_2}{a}\frac{z}{z-1}\left[\frac{1-\epsilon^{-T/\tau_{2m}}}{z-\epsilon^{-T/\tau_{2m}}}\right]\left(\frac{1-\tau_1/\tau_2}{1-{}^{1\tau}/_{2a}}\right)$ $\ldots\left(\frac{1-\tau_{2m-1}/\tau_2}{1-\tau_{2m}/\tau_2}\right)$ $\vdots$ $-\frac{\tau_{2m}}{a}\frac{z}{z-1}\left[\frac{1-\epsilon^{-T/\tau 2m}}{z-\epsilon^{-T/\tau 2m}}\right]\left(\frac{1-\tau_1/\tau_{2m}}{1-{}^{1}/\tau 2m^{a}}\right)$ $\ldots\left(\frac{1-\tau_{2m-1}/\tau_{2m}}{1-\tau_{2m-2}/\tau_{2m}}\right)$	
$\frac{s}{s^2+a^2}$	$\frac{z(z-\cos aT)}{z^2-2z\cos aT+1}$	$\cos(aTn)$
$\frac{a}{(s+\alpha)^2+a^2}$	$\frac{z\epsilon^{-\alpha T}\sin aT}{z^2-2\epsilon^{-\alpha T}z\cos aT+1}$	$\epsilon^{-\alpha Tn}\sin(aTn)$
	$\frac{z}{(z-\epsilon^{-aT})(z-\epsilon^{-bT})}$	$\frac{\epsilon^{-aTn}-\epsilon^{-bTn}}{\epsilon^{-aT}-\epsilon^{-bT}}$
	$\frac{z(a+bz)}{[(z-\alpha)^2+\beta^2](z-1)}$	$\frac{a+b}{(1-\alpha)^2+\beta^2}+\frac{[(a+b\alpha)^2+(b\beta)^2]^{1/2}}{\beta[(\alpha-1)^2+\beta^2]^{1/2}}[\alpha^2+\beta^2]^{n/2}\sin(n\theta+\psi+\lambda)$ $\theta=\tan^{-1}\beta/\alpha,\ \psi=-\tan^{-1}\beta/\alpha-1,\ \lambda=\tan^{-1}\frac{b\beta}{(a+b\alpha)}$
	$\frac{z}{(z-\gamma)[(z-\alpha)^2+\beta^2]}$	$\frac{\gamma^n}{(\gamma-\alpha)^2+\beta^2}+\frac{1}{\beta[(\alpha-\gamma)^2+\beta^2]^{1/2}}(\alpha^2+\beta^2)^{n/2}\sin(n\theta+\psi)$ $\theta=\tan^{-1}\beta/\alpha,\quad \psi=-\tan^{-1}\beta/\alpha-\gamma$
	$\frac{z}{(z-1)(z-\gamma)[(z-\alpha)^2+\beta^2]}$	$\frac{1}{(1-\gamma)[(1-\alpha)^2+\beta^2]}+\frac{[\alpha^2+\beta^2]^{n/2}\sin(n\theta+\Psi+\lambda)}{\beta[(\alpha-\gamma)^2+\beta^2]^{1/2}[(\alpha-1)^2+\beta^2]^{1/2}}$ $+\frac{\gamma^n}{(\gamma-1)[(\gamma-\alpha)^2+\beta^2]}$ $\theta=\tan^{-1}\beta/\alpha,\ \Psi=-\tan^{-1}\beta/\alpha-1,\ \lambda=-\tan^{-1}\beta/\alpha-\gamma$
	$\frac{z(z+a_0)}{(z-1)(z-\gamma)[(z-\alpha)^2+\beta^2]}$	$\frac{(1+a_0)}{(1-\gamma)[(1-\alpha)^2+\beta^2]}+\frac{(\gamma+a_0)^n}{(\gamma-1)[(\gamma-\alpha)^2+\beta^2]}$ $+\frac{[\alpha^2+\beta^2]^{n/2}[(a_0+\alpha)^2+\beta^2]^{1/2}}{\beta[(\alpha-1)^2+\beta^2]^{1/2}[(\alpha-\gamma)^2+\beta^2]^{1/2}}\sin(n\theta+\Psi+\lambda)$ $\theta=\tan^{-1}\beta/\alpha,\ \Psi=\Psi_1+\Psi_2,\ \Psi_1=-\tan^{-1}\beta/\alpha-1$ $\Psi_2=-\tan^{-1}\beta/\alpha-\gamma,\ \lambda=\tan^{-1}\frac{\beta}{(a_0+\alpha)}$
	$\frac{z}{(z-\gamma)(z-1)^2}$	$\frac{\gamma^n}{(\gamma-1)^2}+\frac{1xn}{(1-\gamma)}-\frac{1}{(1-\gamma)^2}$
	$\frac{z(z+a_0)}{(z-\gamma)(z-1)^2}$	$\frac{\gamma+a_0}{(\gamma-1)^2}\gamma^n+\frac{1+a_0}{1-\gamma}n+\left[\frac{1}{1-\gamma}-\frac{a_0+1}{(1-\gamma)^2}\right]$
	$K\frac{z\prod_{i=1}^{m}(z-\mu_i)\prod_{i=1}^{g}[(z-\zeta_i)^2+\gamma_i^2]}{(z-1)\prod_{i=1}^{q}(z-\rho_i)\prod_{i=1}^{l}[(z-\alpha_i)^2+\beta_i^2]}$ $2l+q>2g+m$	$A_0\left[1+\sum_{K=1}^{l}M_K\frac{\beta_0{}^2{}_K[\alpha^2{}_K+\beta^2{}_K]^{n/2}}{\beta_K[(\alpha_K-1)^2+\beta^2{}_K]^{1/2}}\times\right.$ $\left.\sin(n\theta_K+\lambda_K+\Psi_K)+\sum_{K=1}^{q}(-1)^{r_K+K}N_K(\rho_K)^n\right]$

Table I. (Continued)

F(z)	F*(s)	F(nT)
		$A_0=K\dfrac{\prod_{i=1}^{m}(1-\mu_i)\prod_{i=1}^{g}[(1-\xi_i)^2+\gamma_i^2]}{\prod_{i=1}^{e}(1-\rho_i)\prod_{i=1}^{l}[(1-\alpha_i)^2+\beta_i^2]}$ $M_K=\dfrac{\prod_{i=1}^{m}\left(\frac{\mu_{iK}}{1-\mu_i}\right)\prod_{i=1}^{g}\left(\frac{\gamma_{iKa}}{\gamma_{oi}}\right)\left(\frac{\gamma_{iKb}}{\gamma_{oi}}\right)}{\prod_{i=1}^{e}\left(\frac{\rho_{iK}}{1-\rho_i}\right)\prod_{i=1,\,i\neq K}^{l}\left(\frac{\beta_{iKa}}{\beta_{oi}}\right)\left(\frac{\beta_{iKb}}{\beta_{oi}}\right)}$ $N_K=\dfrac{\prod_{i=1}^{m}\left\lvert\frac{\mu_i-\rho_K}{1-\mu_i}\right\rvert\prod_{i=1}^{g}\left(\frac{\bar{\rho}_{Ki}}{\gamma_{oi}}\right)^2}{\prod_{i=1}^{l}\left(\frac{\rho_{Ki}}{\beta_{oi}}\right)^2\prod_{i=1,\,i\neq K}^{e}\left\lvert\frac{\rho_i-\rho_K}{1-\rho_i}\right\rvert}$ $\theta_K=\tan^{-1}\beta_K/\alpha_K$ $\psi_K=\tan^{-1}(\beta_K/\alpha_{K-1})$ $\lambda_K=\sum_{i=1}^{g}\theta_{iK}-\sum_{i=1,\,i\neq K}^{l}\psi_{iK}+\sum_{i=1}^{m}\delta_{iK}-\sum_{i=1}^{e}\phi_{iK}$ r_K = number of real zeros greater than $\lvert\rho_K\rvert$, in the s-plane

MAPPING OF Z-TRANSFORM LOCI IN THE HG*(Z)-PLANE

To discuss the mapping of the loop transfer function $HG^*(z)$ in the $HG^*(z)$-plane as z traverses the unit circle, consider the following cases

1. When $H(s)G(s)$ has simple poles only and no pole at the origin, then $HG^*(z)$ is real both at $\omega=0$ and $\omega=\Pi/T$. Therefore, when z describes one complete revolution along the unit circle, the locus $HG^*(z)$ describes a closed curve in the $HG^*(z)$-plane as shown in Fig. 5.

2. When $H(s)G(s)$ contains simple poles including a pole at the origin, then the corresponding term of $HG^*(z)$ at $z=1$, which is the only effective term, can be written following equation 8, as

$$\underset{s_n=0}{HG^*(z)}=\frac{A(0)}{B'(0)}\,\frac{z}{z-1} \tag{33}$$

Around the point $z=1$, one can write

$$z=1+\rho e^{j\theta} \tag{34}$$

where ρ is an infinitely small radius of a semicircle around $z=1$, and θ is an angle of radius vector ρ, as shown in Fig. 6. Substituting equation 34 in 33, then

$$\lim_{z\to 1} HG^*(z)=\lim_{\rho\to 0}\frac{A(0)}{B'(0)}\times\frac{1+\rho e^{j\theta}}{\rho e^{j\theta}} \tag{35}$$

When ρ describes the small semicircle, as θ varies from $-\Pi/2$, to $\Pi/2$, then $HG^*(z)$ describes an infinite semicircle which varies from $\Pi/2$ to $-\Pi/2$, as shown in Fig. 7.

From the foregoing, only asymptotic shapes of $HG^*(z)$ loci are indicated. However, for the actual plot of $HG^*(z)$ loci in $HG^*(z)$-plane, it is essential to plot each factor in the summation given in equation 8 for $HG^*(z)$ individually. Adding the component loci vectorially one obtains the final shape of $HG^*(z)$ locus. The advantage of this method of plotting z-transform loci lies in the fact that for simple forms of $H(s)G(s)$, each term in the summation indicated in equation 8 is of the following form

$$HG^*(z)=\frac{az}{z+b} \tag{36}$$

where a and b are real or complex. When the locus of z is the unit circle, the bilinear transformation indicated in equation 36 also traverses a circle in the $HG^*(z)$-plane. Thus, $HG^*(z)$ locus is a vector sum of circles, which are very easy to plot. Their magnitude and location are indicative of the final shape of $HG^*(z)$. These constituent circles are determined by the pole, zero configuration of $H(s)G(s)$ in the s-plane. Thus, this leads to synthesis and shaping of $HG^*(z)$ plots by adjusting the value of the poles and zeros of $H(s)G(s)$. The number of computations required in plotting the $HG^*(z)$ loci is identical to that of plotting the $H(s)G(s)$ in the s-plane. Furthermore, the plot of the $HG^*(z)$ locus is exact, so that this method is an improvement over s-domain procedures.

Example

To illustrate the application of the various mathematical expressions discussed or introduced so far, the sampled-data control system shown in Fig. 4 will be analyzed. For this system, the loop transfer function consists of a zero-order hold circuit $G_1(s)$, ordinary dynamical system $G_2(s)$, and unity feedback. (A zero-order hold is a circuit whose function is to convert a single impulse into a rectangular pulse of unit height and T-second duration. The Laplace transform of such a circuit is, as has been derived in reference 5, $G_1(s)=(1-\epsilon^{-Ts})/s$.) The z-transform input-output relation is

$$\theta_c^*(z)=\frac{G_1G_2^*(z)}{1+G_1G_2^*(z)}\,\theta_r^*(z) \tag{37}$$

Suppose

$$G_2(s)=\frac{A}{s(s+a)} \tag{38}$$

then

$$G_1(s)G_2(s)=\frac{(1-\epsilon^{-Ts})A}{s^2(s+a)} \tag{38A}$$

Using Table I the z-transform of $1/s^2(s+a)$ is

$$\frac{(T/a)z}{(z-1)^2}-\frac{(1-\epsilon^{-aT})z}{a^2(z-1)(z-\epsilon^{-aT})} \tag{38B}$$

The z-transform of $G_1(s)G_2(s)$ is $A(1-z^{-1})$ multiplied by equation 38(B). This obtains

$$G_1G_2^*(z)=\frac{AT}{a(z-1)}-\frac{A}{a^2}\,\frac{1-\epsilon^{-aT}}{z-\epsilon^{-aT}} \tag{39}$$

Assume

$$a=1.0,\ A=1.0,\ T=1.0 \tag{40}$$

Substituting equation 39 into equation 37 and inserting values for A, a, T as given in equation 40 obtains

$$\frac{\theta_c^*(z)}{\theta_r^*(z)}=\frac{0.368z+0.264}{z^2-z+0.632} \tag{41}$$

To obtain the response to a step input, take

$$\theta_r(nT)=1 \tag{42}$$

From Table I

$$\theta_r{}^*(z)=\frac{z}{z-1} \tag{43}$$

Substituting equation 43 into 41 gives

$$\theta_c{}^*(z)=\frac{0.368z+0.264}{z^2-z+0.632}\,\frac{z}{z-1} \tag{44}$$

The sampled output is given, applying relation 24, as follows

$$\theta_c(nT)=\frac{1}{2\Pi j}\int_{\Gamma}\frac{0.368z+0.264}{(z-1)(z^2-z+0.632)}z^n dz \tag{45}$$

Γ is a path of integration which encloses the singular points of the integrand in equation 45. The singular points are

$$z=1 \tag{46}$$

$$z=z_1=0.5+j\sqrt{0.382} \tag{47}$$

$$z=z_1=0.5-j\sqrt{0.382} \tag{48}$$

Applying Cauchy's theorem, $\theta_c(nT)$ equals the sum of the residue of the integrand in equation 45. Finally, $\theta_c(nT)$ equals

$$\theta_c(nT)=1-(0.632)^{n/2}\cos n\theta-(0.213)\times(0.632)^{n/2}\sin n\theta \tag{49}$$

where

$$\theta=\tan^{-1}\frac{\sqrt{0.382}}{0.5} \tag{50}$$

The plot of equation 49 for values of n is shown in Fig. 8.

Initial and Final Values

To obtain the initial value of the output equation 26 is applied

$$\theta_c(0)=\lim_{z\to\infty}\frac{z-1}{z}\theta_c{}^*(z)=\lim_{z\to\infty}\frac{0.368z+0.264}{z^2-z+0.632}=0 \tag{51}$$

Application of the final value theorem gives

$$\theta_c(n)_{n\to\infty}=\lim_{z\to 1}\frac{z-1}{z}\,\theta_c{}^*(z)=\lim_{z\to 1}\frac{0.368z+0.264}{z^2-z+0.632}=1 \tag{52}$$

The loop transfer function $G_1G_2{}^*(z)$ as given in equation 39 consists of two terms. The mapping of the first term in $HG^*(z)$-plane as z-traverses the contour shown in Fig. 6 is a straight line (or a circle of infinite radius), as shown in Fig. 9. The locus of the second term is a circle whose center lies on the real axis, as shown in Fig. 9.

The locus of $G_1G_2{}^*(z)$ is obtained by adding vectorially the two constituent terms for each value of z on the unit circle. Since the $G_1G_2{}^*(z)$ locus is symmetric with respect to the real axis, the points corresponding to the upper unit semicircle including 1 and -1 suffice to represent the entire locus. The locus of $G_1G_2{}^*(z)$ is shown in Fig. 9, which indicates that the system is stable for the values of the constants chosen.

Synthesis of Sampled-Data Control Systems

When $HG^*(z)$ has no poles outside the unit circle, the stability of the system is easily assessed from the enclosure or disclosure of the point $(-1,0)$ in the $HG^*(z)$-plane by the locus $HG^*(z)$. Furthermore, the degree of stability of such systems can be improved by shaping the locus of $HG^*(z)$. The subject of main concern in this part is to indicate a means by which shaping is effected. This is done through the insertion of linear networks in the loop of the system. Other means of stabilization and shaping exist, which will be discussed in other papers.

Effect of Simple Poles of H(s)G(s) on HG*(z)

Consider a continuous loop transfer function which consists of simple poles and zeros cascaded with a zero-order hold circuit, given in the following form

$$H(s)G(s)=\frac{A(1-\epsilon^{-Ts})}{s}\times\left[\frac{(1+\tau_1 s)(1+\tau_3 s)(1+\tau_5 s)\ldots(1+\tau_{2m-1}s)}{s(s+a)(1+\tau_2 s)(1+\tau_4 s)\ldots(1+\tau_{2m}s)}\right] \tag{53}$$

where a, τ_1, $\tau_2\ldots\tau_{2m}$ are all real and the degree of the denominator in the brackets of equation 53 is higher than that of the numerator by at least 2. The corresponding z-transform of equation 53 can be easily shown to be

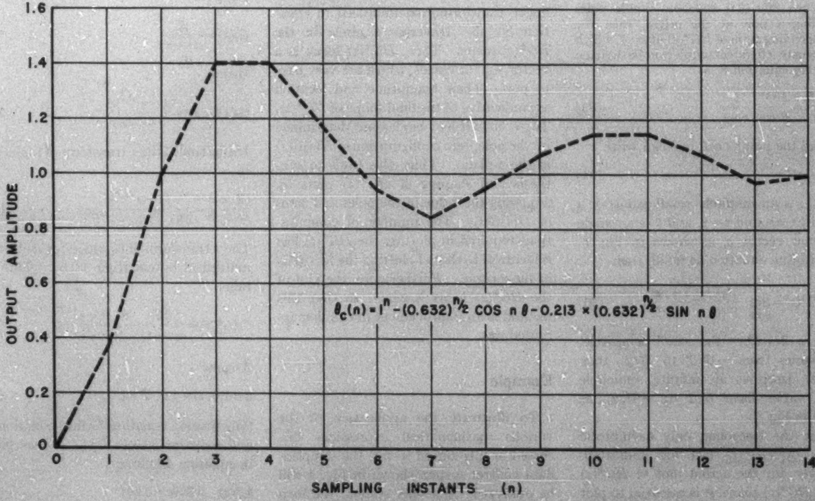


Fig. 8. Plot of output sampled-sequences of system used in example

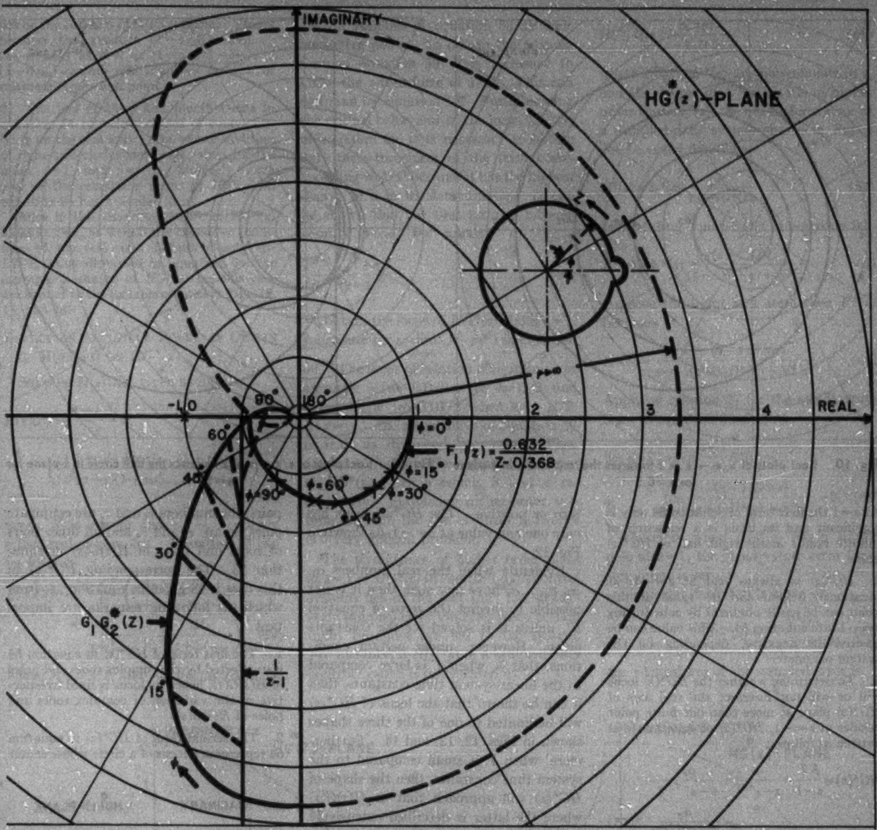


Fig. 9. Locus of $G_1G_2^*(z)$ in example

$$HG^*(z) = \frac{A}{a}\,\frac{T}{z-1} - \frac{A}{a^2}\,\frac{1-\epsilon^{-aT}}{z-\epsilon^{-aT}} \times \frac{(1-a\tau_1)(1-a\tau_3)\ldots(1-a\tau_{2m-1})}{(1-a\tau_2)(1-a\tau_4)\ldots(1-a\tau_{2m})} - \frac{A\tau_2}{a}\left(\frac{1-\epsilon^{-T/\tau_2}}{z-\epsilon^{-T/\tau_2}}\right) \times \frac{(1-\tau_1/\tau_2)(1-\tau_3/\tau_2)(1-\tau_5/\tau_2)\ldots(1-\tau_{2m-1}/\tau_2)}{(1-1/\tau_2 a)(1-\tau_4/\tau_2)(1-\tau_6/\tau_2)\ldots(1-\tau_{2m}/\tau_2)}$$

$$\vdots$$

$$-\frac{A\tau_{2m}}{a}\left(\frac{1-\epsilon^{-T/\tau_{2m}}}{z-\epsilon^{-T/\tau_{2m}}}\right) \times \frac{(1-\tau_1/\tau_{2m})(1-\tau_3/\tau_{2m})\ldots(1-\tau_{2m-1}/\tau_{2m})}{(1-1/\tau_{2m} a)(1-\tau_2/\tau_{2m})\ldots(1-\tau_{2m-2}/\tau_{2m})} \qquad (54)$$

The first term of $HG^*(z)$ in equation 54 is introduced because of a simple pole of $H(s)G(s)$ at the origin (excluding the zero-order hold). The locus of this term is a straight line (or a circle of infinite radius) as has been shown previously. This term is characteristic of sampling systems, for as T increases the locus of $HG^*(z)$ is mainly influenced by this term and consequently the stability of the system is affected. Although T appears in the exponent of other terms, its effect is not substantial. Furthermore, the first term is not affected by the added poles and zeros of $H(s)G(s)$ and consequently independent of the τ's.

The loci of the other terms of $HG^*(z)$ are circles with centers on the real axis of the $HG^*(z)$-plane, in view of the fact that the constants $a, \tau_1, \tau_2 \ldots$ are all real. These circles, while the parameters are varied, are shown in Figs. 10 and 11 for a range of values of each parameter.

Adding vectorially the corresponding points on various circles and the straight line to obtain the final $HG^*(z)$ locus, one can readily observe the effect of each term of $HG^*(z)$, and its correlation to the location of poles and zeros of $H(s)G(s)$. Although the exact shape of $HG^*(z)$ cannot be foreseen before actually plotting $HG^*(z)$ locus point by point, certain characteristics which have significant effects on the final shape of $HG^*(z)$ will be discussed.

1. The locus of $HG^*(z)$ near $z=1$ can be found as indicated previously by introducing a small semicircle of radius ρ around the point $z=1$, and letting ρ approach zero,

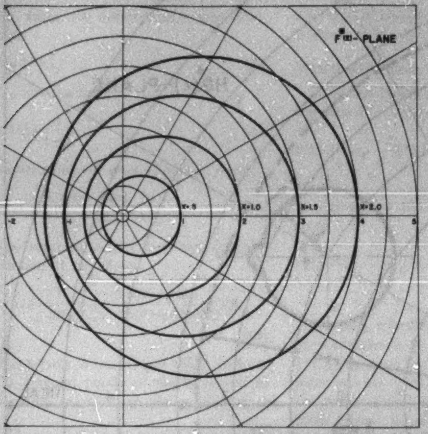


Fig. 10. Loci plots of x/z − a as z traverses the unit circle in z-plane for a = 0.5

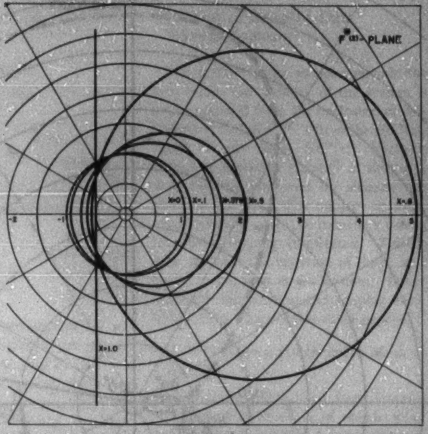


Fig. 11. Loci plots of a/z − x as z traverses the unit circle in z-plane for a = 1

At $z=1$ the first term of equation 54 only is significant and its locus is a semicircle of infinite radius in the right half of $HG^*(z)$ plane.

2. $HG^*(z)$ is always real at $z=-1$, as mentioned before, and its value at this point can be easily obtained by substituting $z=-1$ into equation 54. This value can be positive or negative, depending on the system parameters.

3. To determine whether the $HG^*(z)$ locus will or will not intersect the real axis of $HG^*(z)$ plane at more than one finite point besides at $z=-1$, $HG^*(z)$ of equation 54 is written as follows:

$$HG^*(z)=\frac{KT}{z-1}-\frac{\alpha_1}{z-\epsilon^{-aT}}-\frac{\alpha_2}{z-\epsilon^{-T/\tau_1}}-\dots-\frac{\alpha_n}{z-\epsilon^{-T/\tau_{2m}}} \qquad (55)$$

where $K, \alpha_1, \alpha_2, \dots, \alpha_n$ represent their respective values in equation 54. Since z describes the unit circle in $HG^*(z)$ locus

$$z=\cos\phi+j\sin\phi \qquad (56)$$

Substituting for z in equation 55, and obtaining the imaginary part

$$I_m[HG^*(z)]=\sin\phi\left(\frac{-KT}{2(1-\cos\phi)}+\frac{\alpha_1}{(1+\epsilon^{-2aT})-2\cos\phi\epsilon^{-aT}}+\dots+\frac{\alpha_n}{(1+\epsilon^{-2T/\tau_{2m}})-2\cos\phi\epsilon^{-T/\tau_{2m}}}\right) \qquad (57)$$

equation 57 is zero when $\phi=\Pi$ or $z=-1$, as indicated previously. If the constants $\alpha_1, \alpha_2 \dots \alpha_n$ are all negative, then equation 57 is always negative as ϕ varies from 0 to Π, because all the terms in the brackets will remain negative and never become zero or positive. The $HG^*(z)$ locus has only one real value at $z=-1$, as shown in Fig. 12.

Evidently when the real numbers $\alpha_1, \alpha_2, \alpha_3 \dots \alpha_n$ have any sign, then it is not possible to predict the zeros of equation 57, unless it is solved for the constants given. However, under certain conditions, that is, when T is large compared to the linear-system time constants, then it can be shown that the locus of $HG^*(z)$ will be limited to one of the three shapes shown in Figs. 12, 13, and 14. Furthermore, when T is small compared to the system time constants, then the shape of $HG^*(z)$ will approach that of $H(s)G(s)$ where the latter is described extensively in the literature.

Effect of Complex Roots and Poles of H(s)G(s) on HG*(z)

Since in physical systems complex roots and poles appear in conjugate pairs, the numbers τ_2 and τ_4 are conjugate pairs, so are τ_1 and τ_3 and all other pairs of roots and poles of $H(s)G(s)$ in equation 53. The corresponding $HG^*(z)$ in this case is as given in equation 54, from which the following remarks are important.

1. The first term of $HG^*(z)$ in equation 54 is unaffected by the complex roots and poles of $H(s)G(s)$, hence its locus is fixed irrespective of the variation of complex roots and poles of $H(s)G(s)$.

2. The second term of $HG^*(z)$ in equation 54 represents a locus of a circle whose center

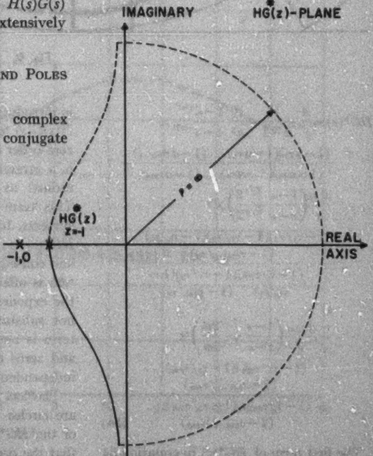


Fig. 12. Typical shape of HG*(z) locus when T is large

lies on the real axis of $HG^*(z)$ plane. This is because the product $(1-a\tau_1)(1-a\tau_3)\ldots(1-a\tau_{2m-1})$ as well as $(1-a\tau_2)(1-a\tau_4)\ldots(1-a\tau_{2m})$ are all real numbers in view of conjugate roots and poles.

3. The loci of third and fourth terms in equation 54 represent circles of center not lying on the real axis. The vector addition of these two circles yield real values for the point $z=1$, and $z=-1$. Similarly, each pair of the remaining terms of $HG^*(z)$ in equation 54 has a similar locus.

Since in this case the closure of $HG^*(z)$ locus at $z=1$, as well as the crossover condition of the real axis, follows closely the $HG(z)$ locus discussed in equation 54, the asymptotic shapes of $HG^*(z)$ for this case are similar to the loci represented in Figs. 12, 13, and 14.

EFFECT OF MULTIPLE POLES AT ORIGIN OF H(s)G(s) ON HG*(z)

Suppose $H(s)G(s)$ given as follows

$$H(s)G(s)=\frac{1-\epsilon^{-Ts}}{s}\times\left(\frac{(1+\tau_1 s)(1+\tau_3 s)\ldots(1+\tau_{2m-1}s)}{s^K(a+s)(1+\tau_2 s)\ldots(1+\tau_{2m}s)}\right)\tag{58}$$

where K is integer and $\tau_1, \tau_2 \ldots \tau_{2m}$ are either real or complex. Equation 58 can be written

$$H(s)G(s)=\frac{1-\epsilon^{-Ts}}{s^n}\left(\frac{A_0}{s+a}+\frac{B_0}{s+1/\tau_2}+\frac{C_0}{s+1/\tau_4}\ldots+\frac{M_0}{s+1/\tau_{2m}}\right)\tag{59}$$

where $n=K+1$, and $A_0, B_0, C_0, D_0 \ldots M_0$ are the residues of equation 58 at their respective poles. To obtain the z-transform of equation 59, it is necessary to know the z-transform of $1/s^n$. This can be found by equation 18. For discussing the critical behaviour of the locus $HG^*(z)$ of equation 59, it is necessary to obtain only the z-transform of the term corresponding to $1/s^n$ because it has the highest degree of s in its denominator. It can be shown that the first term of z-transform of $(1-\epsilon^{-Ts})/s^n$ equals

$$\frac{T^{n-1}}{2}\,\frac{1+z}{(z-1)^{n-1}}\tag{60}$$

To obtain the locus of $HG^*(z)$ about $z=1$, substitute for $z=\lim_{\rho\to 0}(1+\rho\epsilon^{j\theta})$ in equation 60. The locus of relation 60 is infinity at an angle $-(n-1)\Pi/2$ when $\theta=\Pi/2$, and at an angle $(n-1)\Pi/2$ when $\theta=-\Pi/2$. At $z=-1$, the z-transform of equation 58 is real, as discussed previously. Figs. 15 and 16 indicate the asymptotic shapes of $HG^*(z)$ when n equals 3 and 4 respectively. Stabilizing higher-order systems through the use of linear networks is a very difficult problem.

The significance of higher order poles at $z=1$ of $HG^*(z)$ in sampled-data control systems lies in the fact that the steady-state sampled error vanishes when certain input functions are fed to the system. For instance, if $HG^*(z)$ of a sampled-data control system is represented by

$$HG^*(z)=\frac{N^*(z)}{(z-1)D^*(z)}\tag{61}$$

where $N^*(z)$, $D^*(z)$ are polynomials in z, and the degree of $D^*(z)$ in z is of higher order than $N^*(z)$, the z-transform error of a sampled-data system shown in Fig. 4 can be easily derived as[5]

$$E^*(z)=\theta_r^*(z)\,\frac{1}{1+HG^*(z)}\tag{62}$$

Substituting equation 61 in equation 62

$$E^*(z)=\theta_r^*(z)\frac{(z-1)D^*(z)}{(z-1)D^*(z)+N^*(z)}\tag{63}$$

Suppose the input is a step, then $E^*(z)$ becomes

$$E^*(z)=\frac{z}{z-1}\,\frac{(z-1)D^*(z)}{(z-1)D^*(z)+N^*(z)}\tag{64}$$

Applying relation 27 for the steady-state error in equation 64 obtains

$$\lim_{n\to\infty}\epsilon(nT)=\lim_{z\to 1}\frac{z-1}{z}E^*(z)=\lim_{z\to 1}\frac{(z-1)D^*(z)}{(z-1)D^*(z)+N^*(z)}=0\tag{65}$$

Therefore when $HG^*(z)$ has a simple pole at $z=1$, the steady-state error for a step input is zero.

Similarly, if $HG^*(z)$ is given as

$$HG^*(z)=\frac{N^*(z)}{(z-1)^2D^*(z)}\tag{66}$$

the z-transform of the sampled-error is

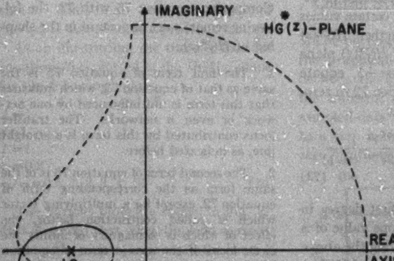


Fig. 13. Typical shape of HG*(z) locus when T is large

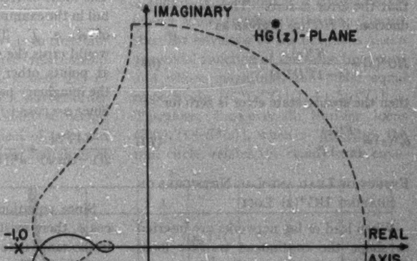


Fig. 14. Typical shape of HG*(z) locus when T is large

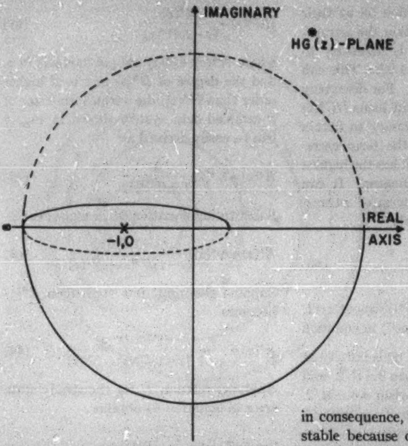


Fig. 15. Typical plot of HG*(z) locus for double pole at origin in H(s)G(s) function without the hold function

$$E^*(z)=\theta_r^*(z)\frac{(z-1)^2D^*(z)}{(z-1)^2D^*(z)+N^*(z)} \tag{67}$$

Suppose the input is a unit ramp

$$\theta_r^*(z)=T\frac{z}{(z-1)^2} \tag{68}$$

Substituting equation 68 in 67, and obtaining the steady-state error, it is noticed that the error is zero. Therefore, by induction, if $HG^*(z)$ is given as

$$HG^*(z)=\frac{N^*(z)}{(z-1)^nD^*(z)} \tag{69}$$

then the steady-state error is zero for

$$\theta_r^*(z)=\frac{F^*(z)}{(z-1)^n} \tag{70}$$

Effect of Lead and Lag Networks on Shaping HG*(z) Loci

When lead or lag networks are inserted to shape the continuous transfer locus $H(s)G(s)$ in continuous control systems, the asymptotic values of the transfer loci do not change for $\omega=0$ and $\omega=\infty$, but the effect of such networks lies primarily in shaping $H(s)G(s)$ between these two frequencies. In contrast to this, the effect of inserting lead or lag networks in sampled-data control systems would change the asymptotic value of $HG^*(z)$ at $z=-1$; in consequence, the entire shape of $HG^*(z)$ is altered. Under certain conditions, the $HG^*(z)$ locus may cross the real axis at one or two points, depending on the values of lead or lag introduced and on the system constants; in consequence, the system might be unstable because of the insertion of either lead or lag networks. The effect of networks on a general sampled-data control system can be deduced from their effect on a second order system† containing a zero-order hold. Suppose $H(s)G(s)$ is

$$H(s)G(s)=\frac{A(1-\epsilon^{-Ts})}{s}\left(\frac{1}{s(s+a)}\right) \tag{71}$$

The corresponding z-transform, utilizing Table I, is

$$HG^*(z)=\frac{AT}{a(z-1)}-\frac{A}{a^2}\frac{1-\epsilon^{-aT}}{z-\epsilon^{-aT}} \tag{72}$$

The locus of $HG^*(z)$ was discussed in detail in the example given for certain values of A, a, T. To test whether $HG^*(z)$ would cross the real axis of $HG^*(z)$ plane at points other than at $z=-1$, equate the imaginary part of equation 72 to zero. Since $z=\cos\phi+j\sin\theta$, then

$$\frac{-AT/a}{2(1-\cos\phi)}+\frac{A(1-\epsilon^{-aT})}{a^2[(1+\epsilon^{-2aT})-2\epsilon^{-aT}\cos\phi]}=0 \tag{73}$$

Since equation 73 is of first degree in $\cos\phi$, there exists at most one value of ϕ in the range $0<\phi<\Pi$, satisfying the above equation. Therefore, the locus of $HG^*(z)$ given in equation 72 may be either one of the shapes shown in Figs. 12 and 13.

A network with transfer function $(1+\tau_1s)/(1+\tau_2s)$, where τ_1 and τ_2 are real, may be lead or lag depending on whether $\tau_1>\tau_2$ or $\tau_1<\tau_2$. Inserting such a typical network into the system described by equation 71, the loop transfer function becomes

$$H(s)G(s)=\frac{A(1-\epsilon^{-Ts})}{s}\times\left(\frac{1}{s(s+a)}\ \frac{1+\tau_1s}{1+\tau_2s}\right) \tag{74}$$

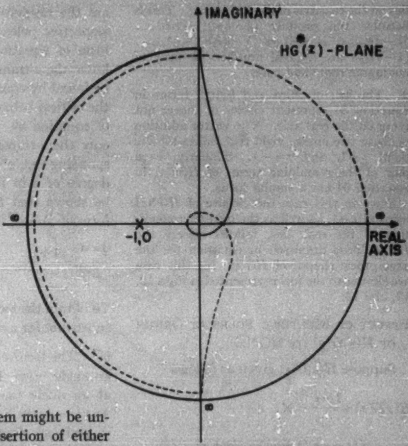


Fig. 16. Typical plot of HG*(z) for triple pole at origin H(s)G(s) function without the hold function

The corresponding z-transform of equation 74 can be obtained using Table I as follows

$$HG^*(z)=\frac{AT}{a(z-1)}-\frac{A}{a^2}\frac{1-\epsilon^{-aT}}{z-\epsilon^{-aT}}\left(\frac{1-\tau_1a}{1-\tau_2a}\right)-\frac{A\tau_2}{a}\frac{1-\epsilon^{-T/\tau_2}}{z-\epsilon^{-T/\tau_2}}\left(\frac{1-\tau_1/\tau_2}{1-1/\tau_2a}\right) \tag{75}$$

Comparing equation 75 with 72, the following remarks are important in the shaping of $HG^*(z)$.

1. The first term of equation 75 is the same as that of equation 72, which indicates that this term is not influenced by one network or even n networks. The transfer locus contributed by this term is a straight line, as indicated before.
2. The second term of equation 75 is of the same form as the corresponding term of equation 72, except for a multiplying factor which is called contraction factor, the effect of which is to magnify or nullify the circle locus of the second term of equation 72, depending on τ_1 and τ_2. In this connection, it is important to mention that only the magnitude of the circle locus of equation 72 is changed, and not its shape.
3. A new term is added by equation 75 the locus of which is a circle whose center is on real axis and whose radius depends on (τ_1/τ_2 and τ_2). This new circular locus is introduced through the pole ($1/\tau_2$) and whose effect is partially to compensate the effect of the circle of the second term. The final shape of $HG^*(z)$ depends on τ_1/τ_2 and τ_2, and by varying these values the shaping

---

† A second-order system is one described by a second-order linear differential equation, i.e., one whose transfer function has a denominator of second degree in s.

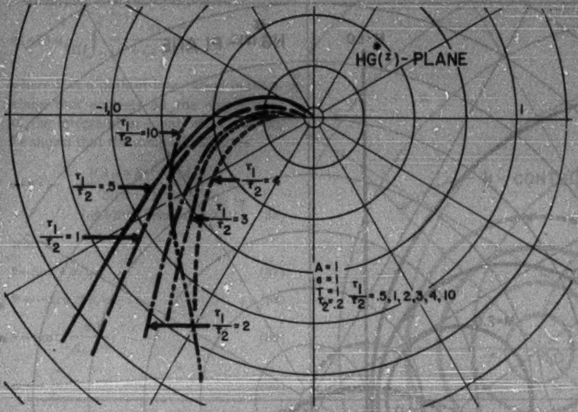

HG(*z)-PLANE
-1,0
1
$\frac{\tau_1}{\tau_2}=10$
$\frac{\tau_1}{\tau_2}=.5$
$\frac{\tau_1}{\tau_2}=4$
$\frac{\tau_1}{\tau_2}=3$
$\frac{\tau_1}{\tau_2}=1$
$\frac{\tau_1}{\tau_2}=2$
A=1
a=1
T=1
$\tau_2=.2$
$\frac{\tau_1}{\tau_2}=.5,1,2,3,4,10$

Fig. 17. Locus plot of HG*(z) in example

of $HG^*(z)$ can be effected. Although shaping is done by trial and error, the shapes of $HG^*(z)$ are limited to those shown in Figs. 12 and 13. In consequence, by plotting few points of the $HG^*(z)$ locus, the final shape can be estimated.

The insertion of n networks of the type $(1+\tau_1 s)/(1+\tau_2 s)$, in equation 71, will introduce n additional terms in $HG^*(z)$, as shown in equation 54. The locus of each term is a circle whose magnitude and shape depend on the network constants. The final expected shape of $HG^*(z)$ is discussed in a previous section.

As an illustration, the transfer locus of equation 75 is plotted in the $HG^*(z)$ plane for the following values

$$A=1$$
$$a=1$$
$$T=1$$
$$\tau_2=0.2$$
$$\tau_1/\tau_2=0.5,\ 1,\ 2,\ 3,\ 4,\ 10 \tag{76}$$

The various plots are shown in Fig. 17. It is observed that when the lead as measured by the ratio τ_1/τ_2 is small the locus of $HG^*(z)$ crosses the real axis at two distinct points. As the lead is further increased, the crossover converges into one point only, that is, when $z=-1$. When further lead is introduced, the crossover point will be nearer the point $(-1,0)$ and the system will then tend to be unstable. Furthermore, when lag is introduced, that is, when $\tau_1/\tau_2=0.5$, the system will also tend to be unstable because $HG^*(z)$ locus will cross the real axis at two distinct points where one of these points will be near the point $(-1,0)$. Therefore, there is an optimum lead that can be applied to this particular system below and above which the system will be less stable. Sampled-data control systems of this type will generally start to oscillate at one of two frequencies, $\omega_r/2$, for $z=-1$, or less than $\omega_r/2$ for complex values of z on the unit circle. The conclusions of this particular example can be generally applied to any type of sampled-data control system when T is large.

Design Criterion for Sampled-Data Control Systems

The purpose of this part is to correlate the graphical-mathematical relationship between the shape of the system loci to the transient response of a second-order sampled-data control system subjected to a step input. The effect of pole and zero locations of the transfer functions of second-order sampled-data control systems on step input response is investigated. Contours of constant gain and phase margins are plotted in $G^*(z)$-plane, whereby shapes of transfer loci can be made to tally with design requirements.

Table II. Relationship Between Time Sequence Response Overshoot and Maximum Frequency Response of Second-Order Sampled-Data Control System

$\omega_0 T$	η	$M_{\max}$	Per Cent Overshoot of Time Sequence Response
1	0.3	4.85	77
1	0.5	1.96	40
1	0.8	1.12	17
1	1.0	1.01	0.18

STEADY-STATE RESPONSE OF SINUSOIDAL INPUT

The significance of $HG^*(z)$ plots is that they relate the steady-state response of the sampled-data control systems to sinusoidal input.[3] It can be shown that if the input is sinusoidal function

$$\theta_r(nT)=\sin(\omega nT) \tag{77}$$

then the steady-state output for unity feedback sampled-data control system is

$$\theta_c(nT)=\left|\frac{G^*(\epsilon^{j\omega T})}{1+G^*(\epsilon^{j\omega T})}\right|\sin(\omega nT+\psi) \tag{78}$$

where ψ the angle extended in the $G^*(z)$-plane, between $G^*(\epsilon^{jT\omega})$ and $1+G^*(\epsilon^{jT\omega})$.

Hence, if the input is a sinusoidal sequence whose envelope is the actual input then the envelope of the output is also a sinusoidal function of the same frequency and whose magnitude and phase equal respectively $|[G^*(\epsilon^{j\omega T})]/[1+G^*(\epsilon^{j\omega T})]|,\psi$. Therefore, the over-all transfer locus $G^*(z)/[1+G^*(z)]$, when z- describes the unit circle, relates the steady-state enve-

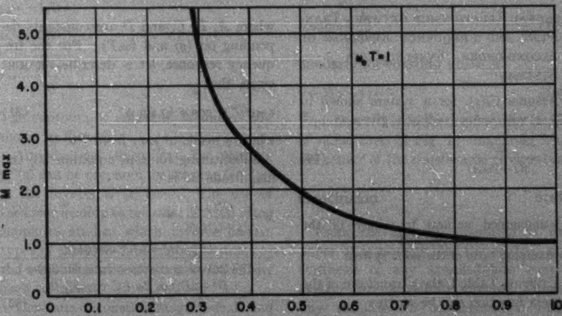


Fig. 18. Plot of $M_{\max}$ versus η for second-order sampled-data control system

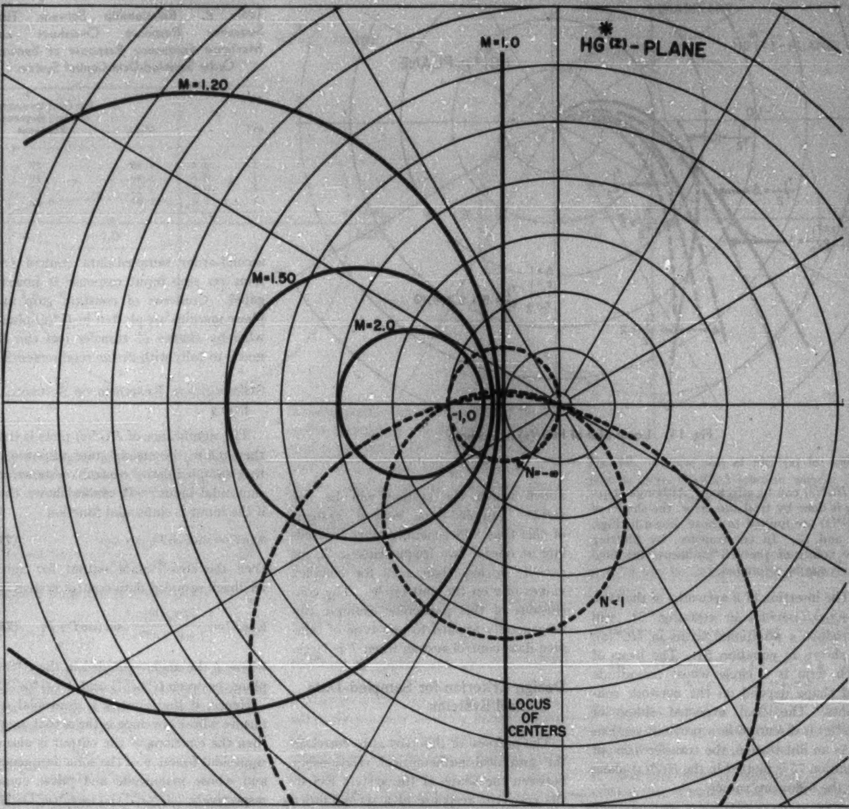


Fig. 19. Constant M-N contours in HG*(z)-plane

lope relationship between input and output.

General Relationship Between Transient and Frequency Response of Second-order Sampled-Data Systems

Assume $G_2(s)$, for a system shown in Fig. 4, with unity feedback, given as

$$G_2(s)=\frac{\omega_0^2}{s(s+2\omega_0\eta)} \tag{79}$$

where

ω_0 = undamped natural frequency of the second-order systems
η = damping ratio of the same system

Utilizing Table I, the z-transform of the system function can be written

$$\frac{G^*(z)}{1+G^*(z)}=\frac{Az+B}{z^2+Cz+D} \tag{80}$$

where A, B, C, and D are constants depending on (η) and (ω_0T). For the frequency response, let z- describe the unit circle, then

$$z=\epsilon^{jT\omega}=\cos\phi+j\sin\phi \tag{81}$$

where ϕ needs to vary from 0 to Π.

Substituting for z in equation 80, the magnitude M is

$$M=\left|\frac{G^*(z)}{1+G^*(z)}\right| =\frac{[B^2+A^2+2AB\cos\phi]^{1/2}}{[C^2+2C(\cos\phi\cos 2\phi+\sin\phi\sin 2\phi)+1+D^2+2D(\cos 2\phi+C\cos\phi)]^{1/2}} \tag{82}$$

Differentiating equation 82, to obtain the angle $\phi_{\max}$ for maximum M gives

$$\cos\phi_{\max}=-\frac{A^2+B^2}{2AB}\pm\sqrt{\left(\frac{A^2+B^2}{2AB}\right)^2-\frac{C(A^2+B^2)(D+1)+2ABD-AB(C^2+D^2+1)}{4ABD}} \tag{83}$$

Substituting equation 83 into equation 82, it is observed that the $M_{\max}$ is a function of A, B, C, D or, alternately, of η and ω_0T.

Having thus obtained the condition for $M_{\max}$, the transient response of the same system can be readily derived. Assuming the input as a step function, then the z-transform of the output can be formed using equation 80 as follows

$$\theta_c^*(z)=\frac{Az+B}{z^2+Cz+D}\,\frac{z}{z-1} \tag{84}$$

The output time-sequence equals

$$\theta_c(nT)=\frac{1}{2\Pi j}\int_\Gamma \frac{Az+B}{(z^2+Cz+D)(z-1)}z^n dz \quad (85)$$

where Γ is a path of integration in the z-plane that encloses all the singularities of the integrand in equation 85. It can be shown that the sampled output is

$$\theta_c(nT)=1+\frac{[(A\alpha+B)^2+(\beta A)^2]^{1/2}}{\beta[(\alpha-1)^2+\beta^2]^{1/2}}\times [\alpha^2+\beta^2]^{n/2}\sin(n\theta+\Psi+\lambda) \quad (86)$$

where

$$\theta=\tan^{-1}\beta/\alpha \quad (87)$$

$$\Psi=-\tan^{-1}\beta/\alpha-1 \quad (88)$$

$$\lambda=\tan^{-1}\frac{\beta A}{A\alpha+B} \quad (89)$$

$$\alpha=-\frac{C}{2} \quad (90)$$

$$\beta=\sqrt{D-\left(\frac{C}{2}\right)^2} \quad (91)$$

Using relation 28, the maximum response can be shown to be

$$\theta_c(\max)=1+\frac{[(A\alpha+B)^2+(\beta A)^2]^{1/2}}{\beta[(\alpha-1)^2+\beta^2]^{1/2}}\times [\alpha^2+\beta^2]^{\frac{\Pi-\lambda}{2\theta}}\sin(\Pi+\Psi) \quad (92)$$

It is observed that equation 92 is also a function of A, B, C, D. Hence, a simple relationship between the maximum of frequency response and the maximum of the time sequence response is established. Table II represents this relationship for different values of η, while ω_0T is held constant. The plot of this table is shown in Fig. 18.

To obtain constant magnitude M and constant phase N contours, write for $G^*(z)$ the following

$$G^*(z)=x^*(z)+jy^*(z) \quad (93)$$

Constant magnitude M can be shown to fulfill the following

$$y^2+\left(x+\frac{M^2}{M^2-1}\right)^2=\frac{M^2}{[M^2-1]^2} \quad (94)$$

Equation is a family of circles in $HG^*(z)$ plane whose centers and radii are given by

$$x_0=-\frac{M^2}{M^2-1} \quad (95)$$

$$y_0=0 \quad (96)$$

$$r=\frac{M}{M^2-1} \quad (97)$$

For constant phase N, write

$$\text{ang}\,\frac{\theta_c^*}{\theta_r^*}(z)=\tan^{-1}y/x-\tan^{-1}y/1+x \quad (98)$$

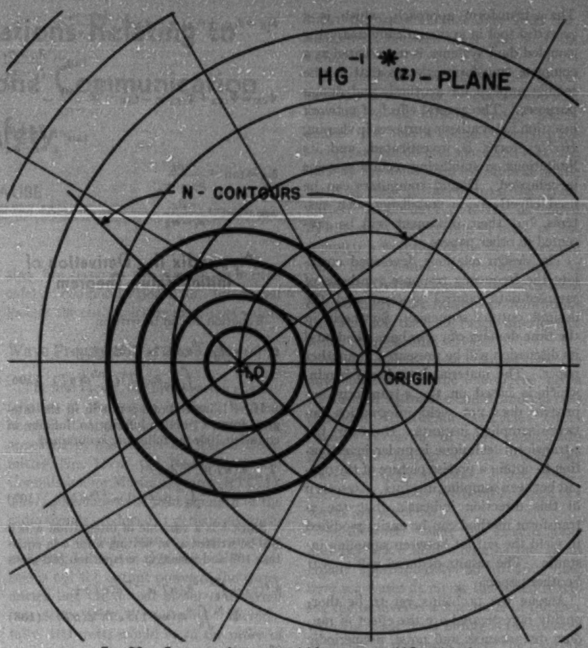


Fig. 20. Contours of constant M-N contours in HG^{-1*}(z)-plane

Putting

$$\tan\text{ang}\,\frac{\theta_c^*}{\theta_r^*}(z)=N \quad (99)$$

equation 98 becomes

$$(x+1/2)^2+\left(y-\frac{1}{2N}\right)^2={}^1/_4\left(\frac{N^2+1}{N^2}\right) \quad (100)$$

which is a family of circles whose centers and radii are given by

$$x_0=-1/2$$

$$y_0=1/2N$$

$$r=\frac{\sqrt{N^2+1}}{2N} \quad (101)$$

The contours of constant M and N are shown in Fig. 19.

An identical treatment for shaping $G^*(z)$ can be prepared for shaping $G^{-1*}(z)$ or the inverse $G^*(z)$-locus. The use of the locus leading to response $G^{-1*}(z)$ often simplifies studies which involve certain types of sampled-data control systems or the use of digital computers in stabilizing such systems.

Contours of constant magnitude M and constant phase N in the inverse plane can readily be deduced following the same procedure used in $G^*(z)$-plane. It can be shown that the constant M curves are of the following form

$$\frac{1}{M^2}=(1+x)^2+y^2 \quad (102)$$

equation 102 is a family of circles (as shown in Fig. 20) of center and radii given by

$$\text{Center}=(-1,0) \quad (103)$$

$$\text{Radius}=\frac{1}{M} \quad (104)$$

Similarly for constant phase N

$$y+N(x+1)=0 \quad (105)$$

Equation 105 is represented by the family of radial lines in Fig. 20 emanating from the center of the constant-M circles.

Conclusion

From the material presented in this paper, it is evident that the analysis and synthesis of any sampled-data control system can be readily effected without any recourse to approximation other than that relating to the narrowness of the pulses that constitute the pulse train.

The z-transform approach, which is a powerful tool in systematically analyzing sampled-data systems, is represented as a complete transform calculus that can be readily applied for synthesis and design purposes. The general effect of network insertion for synthesis purposes in shaping $HG^*(z)$ locus is investigated, and its limitations on stabilizing certain systems is indicated. Digital computers can be used effectively in stabilizing such systems, but their treatment will be presented in other papers.

The design relations developed correlate the frequency response synthesis of sampled-data control systems and the desired output sequence. Synthesis in the time domain can also be applied, and its discussion will be presented in another paper. The material introduced herein, which is based on the z-transform approach, does not yield information between sampling instants. However, the z-transform technique is under investigation to obtain a precise picture of the output between sampling instants. Research in this direction indicates that the z-transform method can be easily modified to yield the output between sampling instants. The results deduced will appear in other papers.

Among the problems yet to be thoroughly investigated are the effect of random disturbances and noise, nonperiodic operation of the sampler, saturation effects, a criterion for obtaining output ripple, consideration of wider sampling pulses, and stabilization methods other than linear networks.

Appendix I. Explanation of z-Transforms and Time Sequence Response Used in Table I

$\alpha_K \pm j\beta_K$ = location of Kth pair of complex poles inside the unit circle
$\zeta_K \pm j\gamma_K$ = location of Kth pair of complex zeros inside the unit circle
ρ_K = location of Kth real pole inside the unit circle
μ_K = location of Kth real zero in the z-plane
$\beta_{0K}^2 = (1-\alpha_K)^2 + \beta_K^2$
$\gamma_{0K}^2 = (1-\zeta_K)^2 + \gamma_K^2$
$\rho_{\iota K}$ = distance from ρ_ι to $\alpha_K + j\beta_K$
$\mu_{\iota K}$ = distance from μ_ι to $\alpha_K + j\beta_K$
$\beta_{\iota Ka}$ = distance from $\alpha_\iota + j\beta_\iota$ to $\alpha_K + j\beta_K$
$\beta_{\iota Kb}$ = distance from $\alpha_\iota - j\beta_\iota$ to $\alpha_K + j\beta_K$
$\gamma_{\iota Ka}$ = distance from $\zeta_\iota + j\gamma_\iota$ to $\alpha_K + j\beta_K$
$\gamma_{\iota Kb}$ = distance from $\zeta_\iota - j\gamma_\iota$ to $\alpha_K + j\beta_K$
$\bar{\rho}_{\iota K}$ = distance from ρ_ι to $\zeta_K + j\gamma_K$

$$\theta_{\iota K} = \theta_{\iota Ka} + \theta_{\iota Kb} = -\tan^{-1}\frac{\beta_K - \gamma_\iota}{\zeta_\iota - \alpha_K} - \tan^{-1}\frac{\beta_K + \gamma_\iota}{\zeta_\iota - \alpha_K}$$

$$\Psi_{\iota K} = \Psi_{\iota Ka} + \Psi_{\iota Kb} = -\tan^{-1}\frac{\beta_K - \beta_\iota}{\alpha_\iota - \alpha_K} - \tan^{-1}\frac{\beta_K + \beta_\iota}{\alpha_\iota - \alpha_K}$$

$$\delta_{\iota K} = \tan^{-1}\frac{-\beta_K}{\mu_\iota - \alpha_K}$$

$$\phi_{\iota K} = \tan^{-1}\frac{-\beta_K}{\rho_\iota - \alpha_K}$$

Appendix II. Derivation of Initial Value Theorem

Equation 25 can be written

$$\theta_c^*(z) = \int_0^{1-0} \theta_c(nT)\epsilon^{-nTs} d(nT) + \int_{1-0}^{\infty} \theta_c(nT)\epsilon^{-nTs} d(nT) \quad (106)$$

If $k+1$, instead of n is put in in the integral from 1 minus infinity to infinity in equation 106, the following is obtained

$$\int_{1-0}^{\infty} \theta_c(nT)\epsilon^{-nTs} d(nT) = \epsilon^{-Ts}\int_0^{\infty} \theta_c(k+1)\epsilon^{-kTs} d(kT) \quad (107)$$

Since k is a variable of integration which can be written as n, writing n for k in equation 107 and inserting in equation 106 gives

$$\theta_c^*(z) = \int_0^{1-0} \theta_c(nT)\epsilon^{-nTs} d(nT) + \epsilon^{-Ts}\int_0^{\infty} \theta_c(n+1)T\epsilon^{-nTs} d(nT) \quad (108)$$

or

$$\int_0^{\infty} \theta_c(n+1)T\epsilon^{-nTs} d(nT) = \epsilon^{Ts}\left(\theta_c^*(z) - \int_0^{1-0} \theta_c(nT)\epsilon^{-nTs} d(nT)\right) \quad (109)$$

Since there exists only one impulse in the interval zero to unity minus zero in equation 109

$$\int_0^{1-0} \theta_c(nT)\epsilon^{-nTs} d(nT) = \int_0^{1-0} \theta_c(nT) d(nT) = \theta_c(0) \quad (110)$$

where $\theta_c(0)$ is the initial value of $\theta_c(nT)$. Furthermore, by definition

$$\mathfrak{L}[\theta_c(n+1)T] = \int_0^{\infty} \theta_c(n+1)T\epsilon^{-nTs} d(nT) \quad (111)$$

Substituting equations 111 and 110 in equation 109 the following is obtained

$$\mathfrak{L}[\theta_c(n+1)T] = \epsilon^{Ts}[\theta_c^*(z) - \theta_c(0)] \quad (112)$$

The Laplace transform of the first difference can be written as

$$\mathfrak{L}[\Delta\theta_c(nT)] = \mathfrak{L}[\theta_c(n+1)T] - \mathfrak{L}[\theta_c(nT)] \quad (113)$$

When equation 113 is substituted in equation 112 and it is noted that $\mathfrak{L}[\theta_c(nT)]$ is $\theta_c^*(z)$, then

$$\mathfrak{L}[\Delta\theta_c(nT)] = z\theta_c^*(z) - z\theta_c(0) - \theta_c^*(z) \quad (114)$$

or

$$z^{-1}\mathfrak{L}[\Delta\theta_c(nT)] = \frac{z-1}{z}\theta_c^*(z) - \theta_c(0) \quad (115)$$

For initial value, let $z \to \infty$, then equation 115 gives the initial value as follows

$$\theta_c(0) = \lim_{z\to\infty} \frac{z-1}{z}\theta_c^*(z) \quad (116)$$

References

1. Sampled-Data Control Systems Studied Through Comparison of Sampling With Amplitude Modulation, **William K. Linvill.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1779–88.

2. Theorie Analytique des Probabilités, Part 1: Du Calcul des Fonctions Generatrices (book), **P. S. Laplace.** Paris, France, 1812.

3. Theory of Servomechanisms (book), **H. M. James, N. B. Nichols, R. S. Phillips.** McGraw-Hill Book Company, Inc., New York, N. Y., chap. 6, 1947.

4. Analyse de Fonctionement des Systems Physique Discontinus, **J. M. Raymond.** *Annales Des Telecommunication*, Paris, France, vol. 4, June, Oct., 1949.

5. The Analysis of Sampled-Data Systems, **J. R. Ragazzini, L. A. Zadeh.** *AIEE Transactions*, vol. 71, pt. II, Nov. 1952, pp. 225–34.

6. Transients in Linear Systems (book), **M. F. Gardner, J. L. Barnes.** John Wiley and Sons, Inc., New York, N. Y., Vol. I, 1942.

7. A List of Generalized Laplace Transforms, **W. M. Stone.** *Journal of Science*, Iowa State College, Ames, Iowa, vol. 22, Apr. 1948.

8. A General Theory of Sampling Servo-Systems, **D. F. Lawden.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 98, pt. IV, Oct. 1951, pp. 31–36.

9. Analysis of a Sampling Servomechanism, **K. S. Miller, R. J. Schwarz.** *Journal of Applied Physics*, New York, N. Y., vol. 21, Apr. 1950, p. 290.

10. Frequency Analysis of Variable Networks, **L. A. Zadeh.** *Proceedings*, Institute of Radio Engineers, New York N. Y., vol. 38, 1950.

11. Fundamental Theory of Servomechanisms (book), **L. A. MacColl.** D. Van Nostrand Company, Inc., New York, N. Y., chap. 10, 1945.

12. A New Approach to the Design of Pulse-Monitored Servo Systems, **A. Porter, F. W. Stoneman.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 97, pt. II, 1950, p. 57.

13. Communications in the Presence of Noise, **Claude E. Shannon.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 37, no. 1, Jan. 1949.

14. An Approximate Transfer Function for the Analysis and Design of Pulsed Servos, **R. G. Brown, G. L. Murphy.** *AIEE Transactions*, vol. 71, pt. II, 1952 (Jan. 1953 section), pp. 435–40.

15. The Dynamics of Automatic Controls (book), **R. C. Oldenbourg, H. Sartorious.** *Mechanical Engineering*, American Society of Mechanical Engineering, New York, N. Y., 1948, chap. 5.

No Discussion

Technical Considerations Relating to Marine Radiotelephone Communication for Safety

O. T. LAUBE
ASSOCIATE MEMBER AIEE

THE International Radio Regulations establishing 2,182 kc as the international calling and distress frequency for maritime mobile radiotelephony in the 1,605- to 2,850-kc band came into effect May 1, 1953.[1] Many government stations have been equipped for operation on this frequency. Parts 7 and 8 of the Rules of the Federal Communications Commission (FCC) require that all ships and public coast stations using telephony in this band be equipped for operation on 2,182 kc. Completion of arrangements for the full utilization of this frequency by ship and coast stations in the United States was scheduled for July 1, 1954.

As a matter of interest to the owners of the more than 30,000 commercial and noncommercial vessels now using 2-megacycle (mc) radiotelephony, it seems appropriate to review the technical factors having a bearing on the performance of this new world-wide safety communication system. It may be useful to begin by examining some of the basic considerations underlying communication, particularly as these relate to the 2 to 3-mc band.

A radiotelephone safety communication system obviously rests upon the ability to conduct conversations between ships, and between ship and coast stations, reliably and consistently. The principal factors involved in making such conversations possible in the 2-mc band are:

1. Propagation characteristics
   a. Ground waves
   b. Sky waves
2. Radio noise
   a. Atmospherics
   b. Man-made sources
3. Interference
   a. Long range
   b. Local
4. Equipment factors
   a. Transmitters
   b. Receivers
   c. Antenna and ground

Most of these factors have been covered in the literature by many authors. Many are matters of continuing study. Since they are basic to the success of marine safety communication, an integrated look at the entire picture may be of value.

Wave Propagation at 2 to 3 Mc

The first factor, the propagation characteristics of the waves in the 2 to 3-mc band, is illustrated in Fig. 1, which first appeared in the report of Special Committee No. 11 of the Radio Technical Commission for Marine Services (RTCM). This figure shows the field strength of the radio signals produced at various distances from a vertical transmitting antenna radiating 100 watts of power. Note that this is not the output power of the transmitter but the actual power transferred into space from the antenna. Incidentally, 100 watts would be in the order of the power radiated at a coast telephone station. Ships would usually fall in the range between 1 and 10 watts. The distances given are in statute miles. The field strength is given in both microvolts per meter (mv/m) and decibels (db) referred to 1 mv/m, a convenient scale for many purposes.

The straight line coming down from the left illustrates the well-known "inverse-distance" law which shows the way in which the radiated energy would drop off with distance due simply to being spread out thinner as the distance increases. This curve, strictly speaking, would hold only for a perfectly conducting plane earth. However, air-ground communications tend to follow this curve for considerable distances.

What actually happens to waves traveling over or near the earth is shown by the three curves to the left of center, those labeled "ground waves." These show what takes place when the path of the waves lies over poor soil, good soil, or fresh water, and over sea water. Ground waves span long distances over areas of high conductivity such as the open seas but are heavily attenuated when low conductivity is encountered, e.g., over terrain which is largely rock, gravel, or dry sand. Propagation over fresh water is about the same as that over good soil, somewhere between poor soil and sea water.

To the right are shown the curves labeled "sky waves." These represent the part of the radiated energy which, unlike the earth-bound ground-wave energy, strikes off into space and is turned back by the ionosphere to reappear on the earth at great distances from the transmitter. In the daytime ionization caused by the sun's rays results in more or less complete absorption of the sky waves and only the ground wave remains. At night sky waves account for occasional exceptionally long-range communications and also for the long-range interference which plagues radio services.

In addition to these changes over the day, conditions in the ionosphere region go through seasonal changes and changes coincident with the sunspot cycle. Other changes occurring at more rapid rates give rise to fading. On the whole, sky waves are quite variable; hence, results are not too readily predictable where sky waves are involved.

At distances where the ground wave and sky wave field strengths are about equal, approximately 50 to 100 miles over sea water at night, the signal reaching the receiver is a composite of the two, with the result that fading is experienced at night, at distances where the daytime signal is quite steady.

Radio Noise in 2 to 3-Mc Band

The second factor, radio noise, embraces two major sources—atmospherics, or noise due to natural phenomena, and the so-called "man-made" noise, usually caused by machinery or electric apparatus—all of which are capable of giving off some energy at radio frequencies. Radio noise enters the receiver along with the desired signal which it tends to obscure. The greater the noise, the greater the signal required to override the noise to provide an adequate margin for intelligibility. Other things remaining equal, this means, of course, that the communication range is reduced as the noise increases.

Atmospheric noise is generally thought

Paper 54-242, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 3, 1954; made available for printing May 6, 1954.

O. T. LAUBE is with the American Telephone and Telegraph Company, New York, N. Y.

The author acknowledges the helpful comments and suggestions received from F. H. Willis of the Bell Telephone Laboratories, Inc., and others.

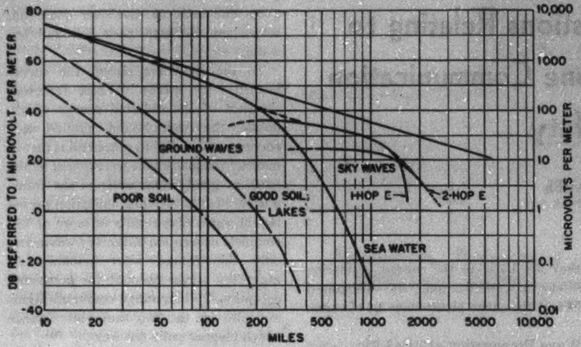


Fig. 1. Two-mc transmission, 100 watts radiated from a vertical antenna

to arise principally from electric storms in the tropical land regions. The radio noise distribution maps, prepared by the Central Radio Propagation Laboratory and showing the noise level contours over the world, illustrate the fact that atmospheric noise varies with latitude, being greatest in the equatorial areas and diminishing toward the poles. It also varies with the season, being greatest in summer and least during the winter.

Atmospheric noise also goes through a daily cycle, illustrated in Fig. 2, which shows how the noise varies from hour to hour throughout the day. This curve has been smoothed for simplicity and represents a median or average of noise values, which actually vary from instant to instant. Instead of plotting the values of the noise as such, this figure has been drawn to present the strength of the radio signal required to override the noise by a sufficient margin to permit conversation. This particular curve applies for 2 mc in the spring season on the northeastern seaboard of the United States.

During the daylight hours atmospheric noise is at a relatively low level. At noon, for example, a signal of only 1 mv/m suffices to permit telephone communication. During the hours after sunset and until sunrise the noise level rises until it requires a signal field some 60 times greater to carry on a conversation. This wide swing in the noise is another manifestation of the effects of changes in propagation of 2-mc radio energy which were discussed earlier. The high noise at night can be attributed to the relatively unattenuated sky wave transmission from the distant tropical areas which are its source. The drop in the daytime results from the extinction of sky wave transmission, leaving only the noise brought in by ground wave.

Man-made noise generally stems from three principal sources: ignition systems of internal-combustion engines, electric machinery and apparatus, and what has sometimes been called "stay" noise or "contact" noise. Noise arising from these causes can be eliminated or, at least, mitigated by various means. Since each db of noise energy must be offset by a db of signal energy by accepting reduced range, measures taken to reduce the noise can be handsomely rewarding in improved performance.

There are two approaches to the problem of noise reduction: one is to locate the receiver and its antenna away from potential sources as far as possible; the other is to suppress the noise at its source. In the case of land stations, where there is usually some freedom of choice, receivers are located at noise-free sites well removed from potential noise sources such as highways, electric railways, factories, and the like. Near-by sources which cannot be avoided are treated to reduce their disturbance to a minimum. By these means the greatest possible reception range is assured. On ships, space is limited and potential noise sources are often numerous. On the smaller craft the main engine with its ignition wiring, an excellent radiator of noise unless adequately shielded, is a major offender. The auxiliaries, motor-driven pumps, fans, generators, compressors, etc., all contribute their share. Even electric razors have been the cause of a sudden rise in the noise level. In fact, any device with a commutator, slip rings, or interrupter, or with contacts which operate frequently, is to some degree a radio noise generator.

Fortunately, most of these sources are amenable to treatment. Small capacitors between motor terminals and ground often are all that is required. Simple resistor-capacitor filters may be needed in some cases. Small radio-frequency "chokes" may also be used. Sometimes noise currents enter the receiver through its power leads. One of the means mentioned, if applied to the leads at the receiver power terminals, will usually take care of such cases. Occasionally, as in the case of ignition wiring, metallic shielding must be provided and connected to ground to confine and by-pass the disturbing radiations.

Stay noise arises principally from the accumulation of electric charges on stays, guys, and other metal parts of the ship's structure which are conducting and either not grounded or poorly grounded. Where these conductors make intermittent contact with grounded parts of the ship, dis-

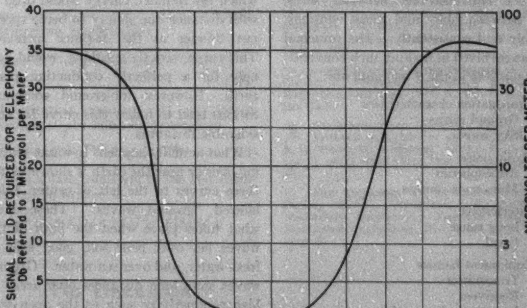


Fig. 2. Hourly variation of atmospheric noise in terms of signal field required for telephony

charges occur which generate noise waves at radio-frequencies. A mast stay, for example, may give rise to noise of this kind because of imperfect electric contact where it passes through the eyebolt which secures it to the hull. Intermittent contact between any metal surfaces at slightly different potentials can produce this effect. On some large vessels where the transmitter is operated continuously during a conversation, stay noise is especially troublesome since the stays and other metal parts of the ship structure are in the strong field of the transmitter. Push button or voice control of the transmitter carrier avoids this condition.

Noise from sources of this general character can usually be mitigated by bonding the parts together, using copper braid or ribbon. Although the corrective measures are simple, the identification and location of the trouble spots is often difficult and time-consuming. A systematic approach, eliminating one possibility after another, will usually yield worth-while improvements.

Interference

The third major factor, interference from other stations, is one which cannot be discussed with quite the assurance used in speaking of the others. It can be a seriously limiting factor; however, the degree to which it will prove limiting remains to be discovered by experience. The reason this is so, of course, is that the use of a common calling and safety frequency on a world-wide basis is a wholly new venture. While the frequency of 2,182 kc has been used for this purpose on the Great Lakes and in Canadian coastal waters for a number of years, the experience gained in these areas is not necessarily indicative of what might be expected when this frequency is put into use all around the country. For one thing, the number of stations using the frequency will be many times greater. Moreover, many more vessels will be using the frequency in the seaboard areas where daytime propagation range is greatest.

Experience in the use of common intership frequencies such as the frequency of 2,738 kc amply demonstrates just how bad conditions can get. However, this is the channel used for intership business and as such it is badly overloaded much of the time. The ordinary use of the 2,182-kc frequency is limited primarily to call and reply purposes; hence it should never be congested with signals as the intership channel is.

Radio waves travel on for great distances beyond the range in which they are useful. There is no way in which these no longer useful waves can be stopped, short of not starting them. Although this sounds rather drastic, it suggests one practical answer which will apply equally well to long-range and short-range interference: limitation of the number and the duration of transmissions. Regulations, both international and domestic, have been drawn up to achieve this end. However, the mere existence of rules is not enough, nor is it enough to leave the matter entirely to the FCC monitoring stations. It is clearly up to the users themselves to avoid creating conditions which will render the channel useless. Each individual user must be judicious in his use of the channel, avoiding all needless talk, if the value of that channel to him and to others is to be preserved. Calling coast stations directly on their working frequencies offers, among other advantages, a reduction in the traffic load and congestion on 2,182 kc.

Long-range interference can also be reduced by reasonable limitation of power. This has been recognized by those who have studied the problem and is reflected in the FCC rules which limit the power to be used on 2,182 kc for transmissions other than distress to 100 watts into the antenna.

In the absence of experience with full-scale operation on 2,182 kc under the conditions prescribed in the rules, quantitative data on the transmission limitations imposed by interference are naturally lacking. However, measurements made in connection with tests of the International Consultative Committee on Radio telephone distress alarm which were conducted on the common intership frequency indicate that these unwanted signals commonly reach levels of 40 mv/m or more. If such levels are encountered on 2,182 kc, its utility will be seriously impaired. Under severe interference conditions a distress call may have to be repeated many times before the channel can be cleared sufficiently to get intelligible messages to those in a position to give aid.

Radio Equipment

The fourth factor, radio equipment, comes down to the things bearing on performance in which the user has some choice. For some ships, notably those compulsorily fitted with radiotelephone equipment under the terms of the Safety of Life at Sea Convention (London, England, 1948) or the Canadian Agreement for the Promotion of Safety on the Great Lakes by Means of Radio (Ottawa Canada, 1951), certain minimum equipment requirements are specified. For the many ships voluntarily equipped the market offers a considerable range of choice to suit the needs of all types of users.

Many equipments are designed particularly for the 2 to 3-mc band; others also include provision for the high frequencies. They range from small compact equipments with transmitters of 5 or 10 watts output power to more elaborate units with transmitters delivering 75 watts or more. Many of the simpler sets are equipped to operate on four different channels. The more elaborate models provide for as many as 10 different channels. Some are arranged for remote control from a small, conveniently located control box. Some provide several receivers to permit listening on several channels at the same time. While most are arranged for "push-to-talk" operation, a few have automatic voice-operated carrier control. It is not possible here to go into much detail about all of the characteristics of the various equipments available. A brief review of those bearing most directly on transmission performance will help in understanding the over-all picture.

Output power of the transmitter has an important bearing on the distance at which signals may be heard. As the 2-mc transmission curve, Fig. 1, shows, over sea water, doubling the transmitter power (increasing by 3 db) adds some 25 to 30 per cent to the range. A tenfold increase in power (10-db increase) would increase the range by two to three times. These figures do not hold under all conditions but afford some measure of the relative performances of transmitters of, say, 5, 10 or 50 watts power.

Ships compulsorily equipped for telephony in accordance with the terms of the Safety of Life at Sea Convention, Regulation 15, are required to have transmitters capable of providing clearly perceptible signals from ship to ship under normal conditions over a distance of at least 150 miles. Clearly perceptible signals are described as signals of at least 25 mv/m. A footnote to the Regulation states that in the absence of measurements it may be assumed that this range will be obtained by 15 watts unmodulated carrier power into an antenna of 25-per-cent efficiency.

In the Agreement for the Promotion of Safety on the Great Lakes by Means of Radio, Regulation 1, a transmitter of 50 watts unmodulated power is specified on ships compulsorily equipped. In this instance it is assumed that with an antenna of 23-per-cent efficiency a fresh-water range of 50 statute miles will be obtained.

Frequency stability, which plays an important part in over-all performance, is no longer troublesome if the equipment and its power source are adequately maintained. All modern transmitters are crystal-controlled and should be well within the tolerances specified by the FCC.

Modulation is also important. Overmodulation is to be avoided because it distorts and garbles the speech and causes "splatter" which spreads out into and interferes with adjacent channels. On the other hand, undermodulation is also to be avoided because it does not use the power of the transmitter effectively. A 100-watt transmitter poorly modulated may have no more effective range than a well-modulated 10-watt transmitter. Unfortunately, however, the 100-watt transmitter, though poorly modulated, loses none of its potential for causing interference.

The FCC has written technical provisions into its rules directed toward assuring proper modulation. Most modern equipment includes some form of limiter circuit to provide a measure of control. However, equipment specifications and designs alone are not enough. It is still necessary to talk into the telephone properly. This fact cannot be emphasized too often since it is one which is frequently responsible for unsuccessful conversations. The rule is simple and applies to any telephone, radio or land line: hold the instrument reasonably close to the mouth and talk directly into it, using a moderate tone of voice. By following this simple rule the user will go a long way toward getting the most out of his equipment.

While the transmitters usually get first mention when ship telephone equipments are discussed, the receivers are of no less importance. Crystal control is an important feature of all good modern marine communications receivers. It greatly simplifies operation and eliminates errors in tuning, thereby assuring rapid, accurate switching as required in modern practice. To be sure of getting the best performance of which a radiotelephone equipment is capable, it is just as important that the receivers be on frequency as it is in the case of the transmitter. Off-frequency operations impose the same penalties, restricted range and poor intelligibility, in both cases. Most good receivers currently available have adequate sensitivity for the reception of signals of about 1 microvolt at the receiver input.

Selectivity is an important factor in receiver design and one which has an important bearing on the performance obtained in a communication system. Data available to special committees of the RTCM which have investigated the matter indicated that, in the 2 to 3-mc band, receivers available in commercial equipments were 25 to 30 db down at the adjacent channel (8 kc removed) and 40 to 55 db down at the alternate channel (16 kc removed). These values of receiver selectivity have been reflected in a recommendation made by the RTCM that the channels immediately adjacent to the 2,182-kc frequency be vacated to provide a guard band around this important frequency. This recommendation has since been embodied in international agreements and in the rules of the FCC. These selectivity data were also used in developing recommendations relating to the allocation of frequencies in the 2 to 3-mc band.

One source of noise which has not been mentioned previously, but which may assume importance under some conditions, resides within the radio receiver itself. This is the noise, sometimes called "set" noise, which is caused by random agitation of electrons in the resistance elements in the receiver input circuits and in its vacuum tubes. Under shipboard receiving conditions this noise is not ordinarily controlling if a proper antenna is provided. At shore stations, on the other hand, where sites may be carefully selected to avoid other noise sources, reception in the higher latitudes may frequently be limited by set noise. Data on good modern receivers with an adequate antenna indicate that, in the 2 to 3-mc band, a signal field in the order of 1 or 2 mv/m will override set noise.

As in the case of transmitters, the power supply contributes importantly to the performance of the receiver. Wide fluctuations in the power supply voltage are often responsible for poor performance and frequent tube failures; hence, the power source should be well regulated. Although data are not readily available for all of the varieties of equipments on the market, it is believed that a deviation of 8 or 10 per cent either way from the nominal supply voltage would be the maximum acceptable. It is preferable that power be provided by means which are inherently stable or which are equipped with automatic voltage-regulating devices.

Shipboard radio equipments are available equipped with two or more radio receivers to permit listening on two or more channels at the same time. The advantages of this arrangement in promoting effective use of frequencies and in improving service have been pointed out by two special committees of the RTCM. Both have recommended the provision of a receiver for each frequency on which a ship expects to receive calls simultaneously. With the implementation of the compulsory watch on the 2,182-kc frequency on all ships using telephony only, the provision of multiple receivers became almost a necessity.

An important part of a radio installation which often receives too little attention is the antenna-ground system. The two must be considered together since it is the two working jointly which provide the means for radiating the power delivered by the transmitter or gather in the signal energy which is utilized by the receiver.

Without going into the reasons why it is so, vertical antennas give best results in a service such as maritime mobile service. All shore and ship stations are therefore provided with vertical antennas. Ideally, an antenna should have physical dimensions which are at least a substantial part of one-half the wave length at the operating frequency. A quarter-wave vertical radiator operating over a good ground makes a very effective radiating system. At 2 mc this means a structure somewhat over 100 feet high.

On shipboard, of course, a radiating system such as this would be out of the question. However, too often the limited possibilities offered on shipboard are not used to the best advantage. In one instance recently, at very little cost, a short, inefficient antenna was replaced by one of better design and a proper ground plate was provided. The improvement obtained was quite astonishing. The power radiated usefully was increased to at least 16 times its original value and the communication range extended to 3 times its former value. When it is considered that even with the improved arrangements only some 25 per cent of the power delivered to the antenna by the transmitter was actually being radiated, the importance of the antenna-ground system becomes even more apparent. At these frequencies the best antennas are none too good and the poorer ones are too wasteful and inefficient to be tolerated.

It is always desirable to make the antenna as high as possible. Where height is limited the antenna may be extended horizontally between masts to provide capacitive loading at the top. In some instances conditions dictate the use of a self-supported vertical whip type of antenna. For greater effectiveness such antennas may be obtained with a coil somewhere near the center of the whip. The coil acts as an inductive center load. In both the capacitive-loaded and inductive-loaded antennas it is the vertical section of the antenna, particularly that between the base and the loading point, which does most of the effective radiating. Loading serves to increase the current flow in the effective vertical portion of the antenna.

The antenna should be located at a point as far as possible from conducting elements such as steel mats, stays, stacks, and other metal parts of the superstructure. Long leads between the transmitter and the base of the antenna cause losses and should be avoided.

A good ground is as important as a good antenna and is not difficult to provide if properly planned at the outset. Losses in the ground can easily become the most important losses in the installation. On vessels with metal hulls a ground wire or braid should be connected directly to the nearest point known to be well grounded to the hull. On vessels with wooden hulls a metal ground plate or strip having an area of at least 10 square feet should be installed on the bottom or alongside the keel.

All of the possibilities cannot be considered in detail here. They are well known to reliable marine radio service people. This brief discussion, however, should make it clear that money spent on the best possible antenna-ground system is a sound investment.

Over-all Performance

As was pointed out at the beginning of the paper, all of the factors examined bear in one way or another on the possibility of communicating by radiotelephone between ships and between ship and shore. These factors include the transmission properties of the frequencies used, the limitations imposed by noise and interference, the capabilities of the equipment, and the manner in which it is installed, maintained, and used.

The effects of all of these factors have been summarized and are shown in Fig. 3, which illustrates the telephone communication ranges to be expected over sea water from hour to hour throughout a day. It has been derived from the propagation data and the atmospheric noise data shown earlier in Figs. 1 and 2 and assumes that the power radiated from the ship's antenna is 10 watts. This figure, incidentally, is representative of the radiated power from a transmitter of, say, 50 watts power working into a good shipboard antenna system.

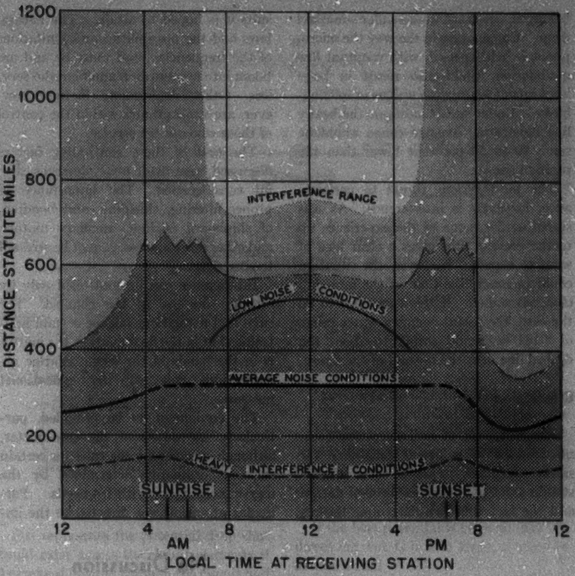


Fig. 3. Distance ranges 2 mc, 10 watts radiated power, spring season, northeastern seaboard of United States

The discrete lines are an attempt to show certain situations which occur often enough to be termed "usual" or "average." The heavy, more or less continuous line is illustrative of the intership ranges which might be expected at times when the channel is quite clear of other conversations, that is to say, when range is limited only by atmospherics and man-made noise on shipboard. At night, when atmospheric noise is at its highest level, it controls the useful range. During the daytime the residual machine noise on shipboard usually sets the limit.

At night, at distances beyond 50 or 100 miles, transmission is by sky wave or a composite of ground and sky waves. During the daylight hours only the ground wave is useful. The line has been broken at sunrise and sunset to indicate marked unpredictability at these periods. During these hours of transition from daylight to darkness, and vice versa, exceptionally long-range conversations are sometimes possible for short periods.

The lighter curve shown in the daytime hours represents the results to be expected if man-made noise were reduced to negligible proportions and atmospherics alone limited transmission. This is the situation which might obtain on transmissions from a ship to a coast station. The difference between the two lines also gives an indication of how much intership range can be improved by clearing up noise sources on shipboard. At times when conditions are particularly favorable the useful range may extend out to the limits of the heavy hatching. During the night hours such results presuppose a predominant sky wave which is fairly steady.

The light hatching illustrates the extent to which the waves travel out beyond their useful range to cause interference at great distances. The nighttime interference range runs off the figure out to some 2,000 or 2,500 miles. Even these distances may be exceeded under some conditions. Incidentally, the ratio of wanted to unwanted signal assumed in arriving at these limits was a rather modest one by most standards.

Thus far it has been assumed that communication is taking place on a more or less cleared channel or one which is not especially busy. It has also been assumed that the equipment is in top condition and is being used correctly. However, if the equipment had not been given proper maintenance attention, or if the power supply voltage was not up where it belonged, the range would suffer a marked drop. Carelessness in the way the microphone is held or used, with resultant low modulation, could also result in large impairment in range as well as in intelligibility. Under such conditions the heavy line illustrating average range would be some 30 to 50 per cent lower than the position shown.

The most serious threat to performance, however, is interference. As was stated earlier, none of the experience up to the present time gives a clear idea of what to expect. However, the situation could be much like that on the intership channels today. If this turns out to be the case, the useful over sea water range of 2,182 kc would shrink to about the dashed line near the bottom of the figure.

Conclusion

The assurance of radiotelephone communications of sufficient reliability for safety purposes requires some understanding of the underlying laws of nature and the factors which determine the results it is hoped to attain. The natural laws and the possibilities and limitations of the frequencies used must be and are taken into account in planning radio services. Other determining factors, however, are substantially within the control of those who use the service.

The first of these controlling factors discussed, man-made noise, deserves careful consideration. The importance of proper filtering, shielding, and bonding, of shipboard engines, auxiliary motors, and other noise sources, cannot be stressed too strongly.

Interference can be controlled only by limiting the use of the channel. The rules and regulations impose certain limitations but in the final analysis the matter is in the hands of the user. Marine organizations can help by educational measures.

The equipment to be provided, particularly the power of the transmitter, although specified by the rules for certain ships, is ordinarily determined by the user in accordance with his needs. Particular attention was directed to the importance of providing a proper antenna and ground, items which are often given inadequate consideration.

The provision of a receiver for each frequency on which a vessel expects to operate in a given area has been mentioned. Multiple receivers offer important possibilities in reducing the use of 2,182 kc for ordinary calls to shore in that, by their use, such calls would be received and completed directly on the working frequency as recommended in the FCC rules. This would not only reduce interference on the calling frequency but also would improve service generally.

Radiotelephony in the 2 to 3-mc band lends itself well to the promotion of safety at sea. Proper attention given to installation, maintenance, and correct use of the radio equipment will assure performance consistent with its potentialities.

Reference

1. Final Acts of the Extraordinary Administrative Radio Conference. Geneva, Switzerland, vol. I, art. 52, 1951.

No Discussion

Forced Oscillations of Nonlinear Circuits

LOUIS A. PIPES
MEMBER AIEE

Synopsis: The method of undetermined coefficients is applied to the solution of typical nonlinear electric circuit problems of importance in electrical engineering. This method is shown to give useful approximate values for the amplitudes and phases of the forced oscillations of the nonlinear circuits considered. The effects of bias potentials on the amplitudes of the oscillations are also determined. The utility of the method is illustrated by applying it to the analysis of forced oscillations of a nonlinear inductor, a saturable reactor, a nonlinear capacitor, and a series-connected magnetic amplifier.

THERE has been considerable activity in recent years in the development of analytical methods for the solution of nonlinear electric circuit problems of practical importance.[1–13] These investigations are based on the assumption that the forcing function applied to the circuit is a sinusoid. There exist many practical circuits such as those involving saturable reactors, saturable capacitors, magnetic and dielectric amplifiers, etc., in which the forcing functions involved contain constant biasing potentials in series with sinusoids. Since the principle of superposition does not apply to nonlinear systems, the required solutions for the circuit response under the action of a forcing function containing a constant plus a sinusoid cannot be readily obtained from the response of the circuit to a forcing function containing a sinusoid alone. It is therefore necessary to develop entirely new solutions for these more complicated forcing functions.

The method presented in this paper is an extension of the method of undetermined coefficients which is a standard method for solving linear problems to the nonlinear case. The method is an approximate one, but gives useful information regarding the phase and amplitude of the forced periodic oscillations. No abstract general theory will be presented but the method will be illustrated by applying it to typical nonlinear circuits of practical importance.

Forced Oscillations of Nonlinear Inductor

An important technical problem is the computation of the amplitude and phase of the fundamental and the harmonic content of the steady-state current in a circuit of the type depicted by Fig. 1.

This circuit contains a harmonic potential $E_m \sin(\omega t+\theta)$ and a bias potential E_0 in series with a linear resistor and a nonlinear inductor. The nonlinear inductor consists of a coil of N turns wound on a magnetic core having a cross-sectional area A and a mean length s.

The differential equation that determines the circuit current i is

---

Paper 54-297, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 15, 1954; made available for printing April 27, 1954.

Louis A. Pipes is with the University of California, Los Angeles, Calif., and the U.S. Naval Ordnance Test Station, Inyokern, China Lake, Calif.

$$Ri+N\dot{\phi}=E_0+E_m \sin (\omega t+\theta) \qquad (1)$$

where ϕ is the magnetic flux in the core of the nonlinear inductor. The flux ϕ is related to the current i by Ampere's law which, expressed in suitable units, is

$$H=Ni/s \qquad (2)$$

where

H = magnetic intensity of the core
s = mean length of the magnetic path in the core

For many practical purposes, the magnetization curve of the core material of the inductor may be represented by the following third-degree polynomial

$$B=\mu_0 H-kH^3 \qquad (3)$$

In this expression B is the magnetic induction and k is a constant determined empirically by adjusting equation 3 to fit the actual magnetization curve of the material of the core. μ_0 is the initial permeability of the core material. It is defined by the equation

$$\mu_0=(dB/dH)_{H=0} \qquad (4)$$

If A is the mean cross-sectional area of the inductor core, the flux ϕ may be expressed in the form

$$\phi=AB=A(\mu_0-kH^3)=A\frac{N}{s}i\mu_0-kA(Ni/s)^3 \qquad (5)$$

Hence

$$N\phi=(AN^2\mu_0/s)i-(kAN^4/s^3)i^3 \qquad (6)$$

It is convenient to write this expression in the following form

$$N\phi=L_0 i-bi^3 \qquad (7)$$

where

$$L_0=\mu_0\frac{N^2A}{s}, \quad b=kAN^4/s^3 \qquad (8)$$

L_0 is the initial inductance of the nonlinear inductor. If $N\phi$ as given by equation 7 is substituted into equation 1 the result is

$$L_0\frac{di}{dt}+Ri-b\frac{d}{dt}i^3=E_0+E_m \sin (\omega t+\theta) \qquad (9)$$

To determine the steady-state response of the nonlinear inductor, it is necessary to determine a periodic solution of equation 9. The method of undetermined coefficients suggests that a periodic solution of the following form be assumed for the current

$$i_0=I_0+I_m \sin (\omega t) \qquad (10)$$

where

I_0 = undetermined d-c component of the steady-state current
I_m = the undetermined amplitude of the fundamental of the steady-state alternating current of the circuit

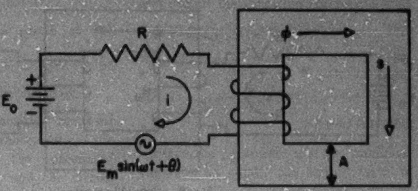


Fig. 1. Nonlinear inductor

To determine I_0 and I_m, substitute equation 10 into the left member of equation 9 and write

$$\begin{aligned}F(t)=&\left(L_0\frac{di_0}{dt}+Ri_0-b\frac{d}{dt}i_0^3\right)\\&=(RI_0)+(RI_m)\sin(\omega t)+\\&\left(\omega L_0 I_m-\frac{3}{4}b\omega I_m^3\omega\right)\cos(\omega t)-\\&3bI_0^2I_m\omega-(3\omega I_0I_m^2b)\sin(2\omega t)+\\&\left(\frac{3}{4}bI_m^3\omega\right)\cos(3\omega t)\end{aligned} \qquad (11)$$

$F(t)$ represents the potential drop that would exist across the circuit elements if the current flowing through the circuit had the form of equation 10. Since the impressed potential of the circuit has the form

$$E(t)=E_0+E_m \sin (\omega t+\theta)=E_0+E_1 \sin (\omega t)+E_2 \cos (\omega t) \qquad (12)$$

where

$$E_1=E_m \cos (\theta), \quad E_2=E_m \sin (\theta) \qquad (13)$$

it is evident that $F(t)\neq E(t)$ and that equation 10 cannot be adjusted[1] to give the exact solution of the differential equation 9. However, an approximate solution of practical utility may be obtained by requiring that the constant term, the sine term, and the cosine term of $F(t)$ be made equal to $E(t)$. This stipulation leads to the following three equations

$$RI_0=E_0 \qquad (14)$$

$$RI_m=E_1=E_m \cos (\theta) \qquad (15)$$

$$\omega L_0 I_m-3b\omega I_m(I_m^2/4+I_0^2)=E_2=E_m \sin (\theta) \qquad (16)$$

These three simultaneous equations serve to determine the unknown amplitudes I_0 and I_m and the phase angle θ between the applied harmonic potential and the fundamental of the resulting alternating current of the circuit. Equation 14 gives the following value for the direct component of the current

$$I_0=\frac{E_0}{R} \qquad (17)$$

The amplitude of the alternating current I_m may be obtained by squaring equations 15 and 16 and adding the result.

This procedure gives

$$I_m^2\left[R^2+\omega^2\left(L_0-\frac{3}{4}bI_m^2-3bI_0^2\right)\right]=E_m^2 \qquad (18)$$

This is a cubic equation in I_m^2 and can be solved by a graphical construction for a given frequency ω. In general equation 18 will have either one or three real roots for the amplitude I_m. The possibility of different amplitudes may lead to "jump phenomena" in special cases.

If the bias potential is large so that the direct current I_0 is also large, it may be assumed that

$$I_m^2/4<<I_0^2 \qquad (19)$$

If the term $I_m^2/4$ is neglected in equation 18, this equation can be solved for the amplitude I_m directly. The result is as follows

$$I_m=\frac{E_m}{\sqrt{[R^2+\omega^2(L_0-3bI_0^2)^2]}} \qquad (20)$$

or

$$I_m=\frac{E_m}{\sqrt{\left[R^2+\omega^2\left(L_0-3\frac{b}{R^2}E_0^2\right)^2\right]}} \qquad (21)$$

Equation 21 shows that the effect of increasing the bias potential is to decrease the effective inductance of the circuit and hence to increase the amplitude of the alternating current. This indicates that by changing the magnitude of the biasing potential it is possible to effect a considerable change in the amplitude I_m of the alternating current of the circuit.

The tangent of the phase angle θ of the alternating current of the nonlinear inductor may be obtained by means of equations 15 and 16 in the form

$$\tan (\theta)=\frac{E_2}{E_1}=\frac{\omega}{R}(L_0-3bE_0^2/R^2) \qquad (22)$$

To this degree of approximation, it is seen that the nonlinear inductor circuit behaves as if it were a linear circuit that has a resistance R and an inductive reactance $X_L=\omega(L_0-3bE_0^2/R^2)$.

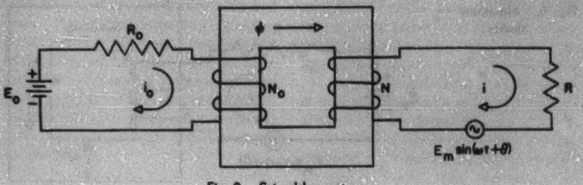


Fig. 2. Saturable reactor

DETERMINATION OF PRINCIPAL HARMONICS

To determine the amplitudes of the principal harmonics of the circuit, write the circuit differential equation 9 in the following form

$$Ri=[E_0+E_1\sin(\omega t)+E_2\cos(\omega t)]-\left(L_0\frac{di}{dt}-b\frac{d}{dt}i^3\right)\quad(23)$$

and substitute $i=I_0+I_m\sin(\omega t)$ into the right member of equation 23. If this is done and relations 14, 15, and 16 are used, the resulting equation may be solved for i. The result is

$$i=I_0+I_m\sin(\omega t)+\left(3\frac{\omega}{R}I_0I_m{}^2b\right)\sin(2\omega t)-\left(\frac{3b}{4R}I_m{}^3\omega\right)\cos(3\omega t)\quad(24)$$

If the higher harmonics are desired, equation 24 may be substituted into 23 and the resulting equation solved for i. The convergence of this procedure requires a separate investigation. If b is a small quantity, it may be shown to converge.

Oscillations of Saturable Reactor

As a second example of the general method, consider the saturable reactor circuit of Fig. 2. This circuit consists of an iron core on which are wound two separate windings. The control winding contains N_0 turns and the output winding N turns. In series with the control windings are a resistance R_0 and a direct potential source E_0. The output winding contains a load resistance R in series with a harmonic potential $E_m\sin(\omega t+\theta)$. By varying the magnitude of the control potential E_0 of the circuit, it is possible to vary the amplitude of the output current.

An application of Kirchoff's voltage law to the control and output circuits yields the following equations

$$R_0i_0+N_0\dot{\phi}=E_0\quad(25)$$

$$Ri+N\dot{\phi}=E_m\sin(\omega t+\theta)\quad(26)$$

where ϕ is the magnetic flux in the iron core. In suitable units, Ampere's circuital law applied to the magnetic circuit of the core leads to the equation

$$N_0i_0+Ni=sH\quad(27)$$

where s is the mean length of the magnetic path of the core. The magnetization curve of the core material may be expressed by the empirical relation

$$H=\frac{\phi}{\mu_0A}+c\phi^3\quad(28)$$

where

H = magnetic intensity
ϕ = core flux
μ_0 = initial permeability of the core material
c = an empirical constant

Equations 25 and 26 may be solved for the currents i_0 and i in the form

$$i_0=\frac{E_0}{R_0}-\frac{N_0}{R_0}\dot{\phi}\quad(29)$$

$$i=\frac{E_m}{R}\sin(\omega t+\theta)-\frac{N}{R}\dot{\phi}\quad(30)$$

If H, i_0, and i are substituted from equations 28, 29, and 30 into equation 27, the result is

$$N_0\left(\frac{E_0}{R_0}-\frac{N_0}{R_0}\dot{\phi}\right)+N\left(\frac{E_m}{R}\sin(\omega t+\theta)-\frac{N}{R}\dot{\phi}\right)=s\left(\frac{\phi}{\mu_0A}+c\phi^3\right)\quad(31)$$

This equation may be put into the more convenient form

$$T\dot{\phi}+\phi+K\phi^3=A+B\sin(\omega t+\theta)\quad(32)$$

where

$$T=\frac{L_0}{R_0}+\frac{L}{R},\ K=\mu_0Ac,\ A=\frac{L_0}{N_0}\frac{E_0}{R_0},\ B=\frac{L_0}{N}\frac{E_m}{R}\quad(33)$$

L_0 and L are two inductances given by

$$L_0=\mu_0\frac{N^2}{s}A,\ L=\mu_0\frac{N^2}{s}A\quad(34)$$

To obtain the steady-state solution of equation 32 by the method of undetermined coefficients, assume a solution of the form

$$\phi=\phi_0+\phi_m\sin(\omega t)\quad(35)$$

Let

$$\begin{aligned}F(t)&=(T\dot{\phi}+\phi+K\phi^3)_{\phi=\phi_0+\phi m\sin(\omega t)}\\&=\left(\phi_0+K\phi_0{}^3+\frac{3}{2}\phi_0\phi_m{}^2\right)+\\&\quad\left[\phi_m+3K\phi_m\left(\frac{\phi}{4}m^2+\phi_0{}^2\right)\right]\times\\&\quad\sin(\omega t)+(T\omega\phi_m)\cos(\omega t)-\\&\quad\left(\frac{K}{4}\phi_m{}^3\right)\sin(3\omega t)-\\&\quad\left(\frac{3K\phi_0}{2}\phi_m{}^2\right)\cos(2\omega t)\end{aligned}\quad(36)$$

If the constant term of $F(t)$, and the coefficients of the sine and cosine terms are equated to the corresponding terms of the right member of equation 32, the following three equations are obtained

$$\phi_0+K\phi_0\left(\phi_0{}^2+\frac{3}{2}\phi_m{}^2\right)=A\quad(37)$$

$$\phi_m+3K\phi_m(\phi_m{}^2+\phi_0{}^2)=B\cos(\theta)\quad(38)$$

$$T\omega\phi_m=B\sin(\theta)\quad(39)$$

These are three simultaneous equations to determine the unknowns ϕ_0, ϕ_m and the phase angle θ. If K is small, the second term of equation 37 may be neglected and the following approximate value for ϕ_0 obtained

$$\phi_0=A=L_0E_0/N_0R_0\quad(40)$$

If the control potential E_0 is large so that $(\phi/4)m^2<<\phi_0{}^2$, the term $(\phi_m{}^2)/4$ may be neglected in the second member of equation 38 and this equation written in the form

$$\phi_m(1+3K\phi_0{}^2)=B\cos(\theta)\quad(41)$$

The sum of the squares of equations 39 and 41 is

$$\phi_m{}^2[T^2\omega^2+(1+3K\phi_0{}^2)^2]=B^2\quad(42)$$

or

$$\phi_m=B/[T^2\omega^2+(1+3K\phi_0{}^2)^2]^{1/2}\quad(43)$$

The tangent of the phase angle θ may be determined by dividing equation 39 by 41. This procedure gives

$$\tan(\theta)=T\omega/(1+3K\phi_0{}^2)\quad(44)$$

If equations 40 and 43 are substituted into equation 35, the following value for the core flux ϕ is obtained

$$\phi=\frac{L_0}{N_0}\frac{E_0}{R_0}+LE_m\sin(\omega t)/NR[T^2\omega^2+(1+3K\phi_0{}^2)^2]^{1/2}\quad(45)$$

This expression gives the approximate value of the core flux to terms of fundamental order.

DETERMINATION OF PRINCIPAL HARMONICS OF CORE FLUX

To determine the principal harmonics of the core flux, write equation 32 as

$$\phi=A+B\sin(\omega t+\theta)-(T\dot{\phi}+K\phi^3) \qquad (46)$$

If $\phi=\phi_0+\phi_m\sin(\omega t+\theta)$ is substituted into the right member of equations 46 and 37, equations 38 and 39 are used to simplify the resulting expression and the following equation is obtained

$$\phi=\phi_0+\phi_m\sin(\omega t)+\frac{K}{4}\phi_m^3\sin(3\omega t)+\frac{3}{2}K\phi_0\phi_m^2\cos(2\omega t) \qquad (47)$$

The control current i_0 may now be determined by equation 29 in the form

$$\begin{aligned}i_0&=\frac{E_0}{R_0}-\frac{N_0}{R_0}\dot{\phi}\\&=\frac{E_0}{R_0}-\frac{N_0}{R_0}\omega\phi_m\left(\cos(\omega t)+\frac{3}{4}K\phi_m^2\times\cos(3\omega t)-3K\phi_0\phi_m\sin(2\omega t)\right)\end{aligned} \qquad (48)$$

The output current i can be determined by the use of equation 30 and expressed in the following form

$$\begin{aligned}i&=\frac{E_m}{R}\sin(\omega t+\theta)-\frac{N}{R}\dot{\phi}\\&=\frac{E_m}{R}\sin(\omega t+\theta)-\frac{N}{R}\omega\phi_m\times\left(\cos(\omega t)+\frac{3}{4}K\phi_m^2\cos(3\omega t)-3K\phi_0\phi_m\sin(2\omega t)\right)\end{aligned} \qquad (49)$$

The effect of the control potential E_0 in influencing the amplitude of the output current is clearly apparent if ϕ_0 as given by equation 40 is substituted into equation 43 and ϕ_m written in the following form

$$\phi_m=\frac{LE_m}{NR}\left[T^2\omega^2+(1+3KL_0^2E_0^2/N_0^2R_0^2)\right]^{-1/2} \qquad (50)$$

It is evident from equation 50 that an increase in the control potential E_0 produces a decrease in ϕ_m and this in turn produces a change in the output current i as is apparent from equation 49.

Forced Oscillations of Nonlinear Capacitor

As a third example of the general procedure, consider the circuit of Fig. 3. This circuit consists of a linear resistor and inductor in series with a nonlinear capacitor. The circuit is energized by a harmonic potential $E_m\sin(\omega t+\theta)$ and a direct biasing potential E_0. It will be assumed that the saturation curve of the nonlinear capacitor may be approximated by the following cubic polynomial[9,10]

$$E_c=S_0q+aq^3 \qquad (51)$$

where

E_c = potential across the plates of the nonlinear capacitor

q = charge separation on the plates of the capacitor

S_0 = initial elastance of the capacitor

The coefficient a is a positive constant that depends on the characteristics of the nonlinear dielectric of the capacitor.[9] The capacitor charge q and the circuit current i satisfy the differential equation

$$L\frac{di}{dt}+Ri+S_0q+aq^3=E_0+E_m\sin(\omega t+\theta) \qquad (52)$$

To obtain the steady-state periodic solution, the method of undetermined coefficients suggests that a periodic solution of the following form be assumed for the current and charge of the circuit

$$i=I_m\sin(\omega t) \qquad (53)$$

$$q=Q_0-\frac{I_m}{\omega}\cos(\omega t) \qquad (54)$$

Q_0 is the constant charge accumulation on the plates of the capacitor and I_m is the maximum amplitude of the alternating current of the circuit. We now substitute equations 53 and 54 into the left member of equation 52 and let

$$\begin{aligned}F(t)&=L\left(\frac{di}{dt}+Ri+S_0q+aq^3\right)_{i=I_m\sin(\omega t)}\\&=[S_0Q_0+aQ_0^3+3aQ_0I_m^2/2\omega^2]+[RI_m]\sin(\omega t)+\left(\omega LI_m-S_0\frac{I_m}{\omega}-3aI_m^3/4\omega^3-3aQ_0^2I_m/\omega\right)\cos(\omega t)-\\&\quad[aI_m^3/4\omega^3]\cos(3\omega t)+3aQ_0I_m^2\cos(2\omega t)/2\omega^2\end{aligned} \qquad (55)$$

Following the general procedure, we now equate the coefficients of the constant, sin (ωt) and cos (ωt) terms of $F(t)$ and the corresponding terms of

$$E(t)=E_0+E_m\sin(\omega t+\theta)=E_0+E_m\cos(\theta)\sin(\omega t)+E_m\sin(\theta)\cos(\omega t) \qquad (56)$$

and obtain the following three equations

$$S_0Q_0+aQ_0(Q_0^2+3I_m^2/2\omega^2)=E_0 \qquad (57)$$

$$RI_m=E_m\cos(\theta) \qquad (58)$$

$$I_m(\omega L-S_0/\omega)-3\frac{a}{\omega}I_m(Q_0^2+I_m^2/4\omega^2)=E_m\sin(\theta) \qquad (59)$$

Since the parameter a of the nonlinear capacitor is small, the second term of equation 57 may be neglected and the following approximate value obtained for Q_0

$$Q_0=\frac{E_0}{S_0}=C_0E_0 \qquad (60)$$

where $C_0=1/S_0$ is the initial capacitance. If the bias potential E_0 is large in comparison with the maximum value of the applied harmonic potential so that $E_0>>E_m$, then for the usual frequencies used in practice we have

$$Q_0^2>>I_m^2/4\omega^2 \qquad (61)$$

If the term $I_m^2/4\omega^2$ is neglected in equation 59, this equation may be written in the form

$$I_m\left(\omega L-\frac{1}{\omega}(S_0+3aC_0^2E_0^2)\right)=E_m\sin(\theta) \qquad (62)$$

Hence if equations 58 and 62 are squared and the results added, we obtain the following equation

$$I_m^2\left[R^2+\left(\omega L-\frac{1}{\omega}(S_0+3aC_0^2E_0^2)\right)^2\right]=E_m^2 \qquad (63)$$

The amplitude of the alternating current of the circuit is therefore given by

$$I_m=E_m/\left[R^2+\left(\omega L-\frac{1}{\omega}(S_0+3aC_0^2E_0^2)\right)^2\right]^{1/2} \qquad (64)$$

It is thus apparent that the effect of the bias potential is to increase the effective elastance of the system by an amount $3aC_0^2E_0^2$.

The tangent of the phase angle θ may

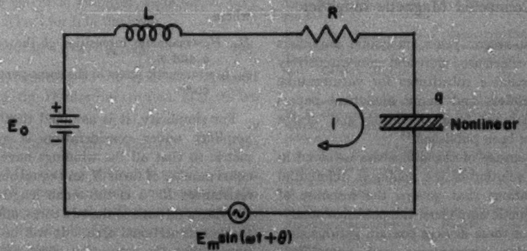


Fig. 3. Nonlinear capacitor

be obtained by the division of equation 62 by 52. The result is

$$\tan(\theta)=\left(\omega L-\frac{1}{\omega}(S_0+3aC_0^2E_0^2)\right)/R \qquad (65)$$

It may be noticed that if the circuit of Fig. 3 does not contain inductance then $L=0$, and equation 65 reduces to

$$I_m=E_m/[R^2+(S_0+3aC_0^2E_0^2)^2/\omega^2]^{1/2} \qquad (66)$$

The effect of the bias potential in increasing the effective impedance of the circuit is clearly evident in equation 66. It is thus apparent than an increase of the magnitude of the bias potential produces a corresponding decrease in the amplitude of the alternating current. This principle is the basis of the operation of dielectric amplifiers.[9]

Determination of Principal Harmonics

The amplitudes of the principal harmonics of the circuit may be determined by writing equation 52 in the following form

$$S_0q=[E_0+E_m\sin(\omega t+\theta)]-\left(L\frac{di}{dt}+Ri+aq^3\right) \qquad (67)$$

and substituting equations 53 and 54 into the right member of equation 67.

If equations 57, 58 and 59 are used to simplify the resulting expression, it can be written in the form

$$q=Q_0-\frac{I_m}{\omega}\cos(\omega t)+aI_m^3\cos(3\omega t)/4\omega^2S_0-3aQ_0I_m^2\cos(2\omega t)/2\omega^2S_0 \qquad (68)$$

The circuit current is then given by

$$i=\frac{dq}{dt}=I_m\sin(\omega t)-3aC_0I_m^3\sin(3\omega t)/4\omega^2+3aQ_0I_m^2\sin(2\omega t)/\omega S_0 \qquad (69)$$

The current therefore contains a second- and third-harmonic current component.

Steady-State Oscillations of Series-Connected Magnetic Amplifier

In recent years, magnetic amplifiers have been used more and more extensively as reliable substitutes for vacuum-tube amplifiers and many analytical papers explaining their operation and design have been published.[13]

Because of the difficulties involved in the solution of the nonlinear differential equations that govern the response of magnetic amplifiers, the published analyses of these devices contain various simplifying assumptions and approximations.[14] The method of undetermined coefficients gives useful information when applied to the study of the response of magnetic amplifiers. To illustrate the general principles involved, the method will be applied to the study of the steady-state response of the symmetrical series-connected magnetic amplifier depicted in Fig. 4.

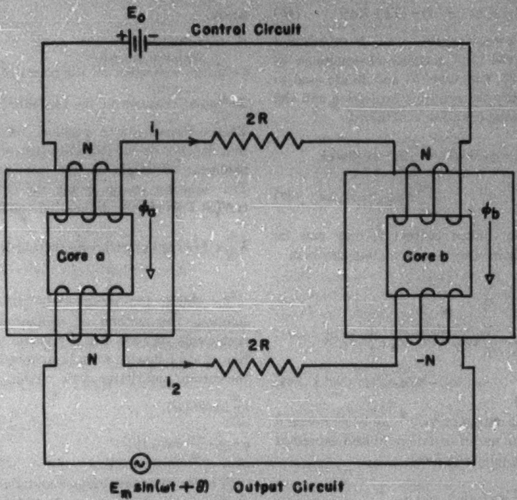


Fig. 4. Series-connected magnetic amplifier

The basic equations of the circuit of Fig. 4 are obtained by an application of Kirchhoff's voltage laws and Ampere's circuital law for the electric and magnetic loops of the amplifier. These equations are

$$2Ri_1+N(\phi_a+\phi_b)=E_0 \qquad (70)$$

$$2Ri_2+N(\phi_a-\phi_b)=E_m\sin(\omega t+\theta) \qquad (71)$$

$$N(i_1+i_2)=sH_a \qquad (72)$$

$$N(i_1-i_2)=sH_b \qquad (73)$$

where

H_a, H_b = magnetic intensities of the cores a and b
ϕ_a, ϕ_b = magnetic fluxes in the corresponding cores

For simplicity, it is assumed that the amplifier under consideration is symmetric so that all the windings have an equal number of turns N, and equal ohmic resistances R. s is the mean length of the magnetic path of the two cores, and A their cross-sectional area. It will be assumed that the saturation curve of the core material is such that the relation between the magnetic intensity and the magnetic flux of the cores may be expressed by the following equations

$$H_a=\phi_a/\mu_0A+c\phi^3 \qquad (74)$$

$$H_b=\phi_b/\mu_0A+c\phi^3 \qquad (75)$$

where

μ_0 = initial permeability of the core material
c = an empirical constant

If equations 70 and 71 are added and the sum (i_1+i_2) is expressed in terms of H_a by equation 72, the result is

$$\phi_a+RsH_a/N^2=[E_0+E_m\sin(\omega t+\theta)]/2N \qquad (76)$$

If equation 71 is subtracted from 70 and (i_1-i_2) is expressed in terms of H_b by equation 73, the result is

$$\phi_b+RsH_b/N^2=[E_0-E_m\sin(\omega t+\theta)]/2N \qquad (77)$$

H_a and H_b may be expressed in terms of ϕ_a and ϕ_b by equations 74 and 75, and equations 76 and 77 may be written in the following form

$$\phi_a+R\phi_a/L_0+k\phi_a^3=[E_0+E_m\sin\times(\omega t+\theta)]/2N \qquad (78)$$

$$\phi_b+R\phi_b/L_0+k\phi_b^3=[E_0-E_m\sin\times(\omega t+\theta)]/2N \qquad (79)$$

where

$$L_0=\mu_0N^2A/s,\quad k=Rsc/N^2 \qquad (80)$$

The quantity L_0 is the initial inductance of one of the core windings. Equations 78 and 79 for the fluxes of the cores are both of the general form given in the following

$$\dot{\phi}+R\phi/L_0+k\phi^3=[E_0+E_m \sin(\omega t+\theta)]/2N \tag{81}$$

To obtain the approximate steady-state solution of this equation by the method of undetermined coefficients, assume a solution of the form

$$\phi=\phi_0+\phi_m \sin(\omega t) \tag{82}$$

If this expression for ϕ is substituted into the left member of equation 81, the result is

$$\begin{aligned} F(t)=(\dot{\phi}+R\phi/L_0+k\phi^3)_{\phi=\text{[illegible]}\sin(\omega t)} \\ =\left[\frac{R}{L_0}\phi_0+k\phi_0\left(\phi_0^2+\frac{\phi_m^2}{2}\right)\right]+ \\ \left[\frac{R}{L_0}\phi_m+3k\phi_m\left(\frac{\phi_m^2}{4}+\phi_0^2\right)\right]\times \\ \sin(\omega t)+(\omega\phi_m)\cos(\omega t)- \\ \left(\frac{k}{2}\phi_0\phi_m^2\right)\cos(2\omega t)- \\ \left(\frac{k}{4}\phi_m^3\right)\sin(3\omega t) \end{aligned} \tag{83}$$

The constant term and the coefficients of sin (ωt) and cos (ωt) are now equated to the corresponding terms of the right member of equation 81. This procedure leads to the following three equations

$$\frac{R}{L_0}\phi_0+k\phi_0\left(\phi_0^2+\frac{\phi_m^2}{2}\right)=E_0/2N \tag{84}$$

$$\frac{R}{L_0}\phi_m+3k\phi_m\left(\frac{\phi_m^2}{4}+\phi_0^2\right)=E_m\cos(\theta)/2N \tag{85}$$

$$\omega\phi_m=E_m\sin(\theta)/2N \tag{86}$$

The system of algebraic equations 84, 85, and 86 determines the unknown quantities ϕ_0, ϕ_m, and the phase angle θ. The solution of this system of equations presents formidable difficulties but an approximate solution may be effected by realizing that the constant $k=Rsc/N^3$ is a small quantity in cases of practical importance. If the term containing k in equation 84 is neglected in comparison with the first term, then this equation may be solved for ϕ_0 with the following result

$$\phi_0=L_0E_0/2NR \tag{87}$$

Now, for sufficiently large values of the bias potential E_0, the assumption $\phi_m^2 << \phi_0^2$ may be made, and equation 85 may be written in the form

$$\phi_m\left(\frac{R}{L_0}+3k\phi_0\right)=E_m\cos(\theta)/2N \tag{88}$$

If equations 86 and 88 are squared and the results added, the following equation is obtained

$$\phi_m^2[\omega^2+(3k\phi_0^2+R/L_0)^2]=E_m^2/4N^2 \tag{89}$$

or

$$\phi_m=E_m/2N[\omega^2+(R/L_0+3k\phi_0^2)^2]^{1/2} \tag{90}$$

Hence the approximate solution of equation 81 is given by

$$\phi=L_0E_0/2NR+E_m\sin(\omega t)/2N[\omega^2+(R/L_0+3k\phi_0^2)^2]^{1/2} \tag{91}$$

The tangent of the phase angle θ is obtained by dividing equation 86 by 88 and obtaining

$$\tan(\theta)=\omega/(R/L_0+3k\phi_0^2) \tag{92}$$

PRINCIPAL HARMONICS OF CORE FLUX

To compute the principal harmonics of the core flux, write equation 81 in the following form

$$R\phi/L_0=[E_0+E_m\sin(\omega t+\theta)]/2N-(\dot{\phi}+k\phi^3) \tag{93}$$

and substitute $\phi=\phi_0+\phi_m\sin(\omega t)$ in the right member of equation 93. If this is done and equations 84, 85, and 86 are used to simplify the result, the following expression for ϕ is obtained

$$\phi=\phi_0+\phi_m\sin(\omega t)+(kL_0\phi_0\phi_m^2/2R)\times \cos(2\omega t)+\left(\frac{k}{4R}L_0\phi_m^3\right)\sin(3\omega t) \tag{94}$$

To this degree of approximation, the core flux is seen to have second- and third-harmonic components. Comparing equations 78 and 79 for the determination of ϕ_a and ϕ_b with equation 81, it is seen that the fluxes ϕ_a and ϕ_b are, in the periodic steady state

$$\phi_a=\phi_0+\phi_m\sin(\omega t)+A_2\cos(2\omega t)+B_3\sin(3\omega t) \tag{95}$$

$$\phi_b=\phi_0-\phi_m\sin(\omega t)+A_2\cos(2\omega t)-B_3\sin(3\omega t) \tag{96}$$

where

$$A_2=(kL_0\phi_0\phi_m^2/2R),\ B_3=(kL_0\phi_m^3/4R) \tag{97}$$

To determine the control current i_1 of the magnetic amplifier, equation 70 may be solved for i_1 in the form

$$i_1=E_0/2R-N(\dot{\phi}_a+\dot{\phi}_b)/2R \tag{98}$$

If equations 95 and 96 are differentiated and the results substituted into 98 we obtain

$$i_1=E_0/2R+2N\omega A_2\sin(2\omega t)/R \tag{99}$$

and the control current is seen to have a small second-harmonic component of the output potential. The output current i_2 may be obtained by solving equation 71 in the form

$$i_2=E_m\sin(\omega t+\theta)/2R+N(\dot{\phi}_b-\dot{\phi}_a)/2R \tag{100}$$

If the time derivatives of equations 95 and 96 are substituted into 100 the result is

$$i_2=E_m\sin(\omega t+\theta)/2R-\omega N[\phi_m\cos(\omega t)+3B_3\cos(3\omega t)]/R \tag{101}$$

and the output current is seen to contain a third-harmonic component of the applied potential. The effect of the control potential on the output current may be clearly seen if equation 87 is substituted into 90 and ϕ_m written in the following form

$$\phi_m=E_m/2N[\omega^2+(R/L_0+3kL_0^2E_0^2/4N^2R^2)^2]^{1/2} \tag{102}$$

Conclusions

The analysis presented in this paper is an attempt to apply the classical method of undetermined coefficients which has proved so useful in the analysis of linear systems so as to obtain the approximate solutions of the basic differential equations of typical nonlinear electric circuits of importance in engineering.

The examples of the application of the method indicate that useful practical results may be obtained by this procedure in a very simple manner. The method appears to be particularly well adapted to the study of electric circuits that contain biasing control potentials. The effect of the control potentials on the current amplitudes involved and the amplitudes of the principal harmonics that are present in the circuits may be easily determined by the method of undetermined coefficients.

References

1. NONLINEAR THEORY OF ELECTRIC OSCILLATIONS, B. van der Pol. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 22, 1934, pp. 1051–86.

2. FORCED OSCILLATIONS IN NEARLY SINUSOIDAL SYSTEMS, M. L. Cartwright. *Journal*, Institution of Electrical Engineers, London, England, vol. 95, pt. III, 1948, pp. 88–96.

3. THE APPLICATION OF POWER SERIES TO THE SOLUTION OF NONLINEAR CIRCUIT PROBLEMS, A. W. Gillies. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 96, pt. III, 1949, pp. 453–75.

4. FORCED OSCILLATIONS IN OSCILLATOR CIRCUITS AND THE SYNCHRONIZATION OF OSCILLATORS, D. G. Tucker. *Journal*, Institution of Electrical Engineers, London, England, vol. 92, pt. III, 1946, p. 226.

5. FORCED OSCILLATIONS IN NONLINEAR SYSTEMS, M. L. Cartwright. *Journal of Research*, National Bureau of Standards, Washington, D. C., vol. 45, 1950, p. 514.

6. THE REVERSION METHOD OF SOLVING NONLINEAR DIFFERENTIAL EQUATIONS, L. A. Pipes. *Journal of Applied Physics*, New York, N. Y., vol. 23, 1952, p. 202.

7. STABILITY INVESTIGATIONS OF THE NONLINEAR PERIODIC OSCILLATIONS, C. Hayashi. *Journal of Applied Physics*, New York, N. Y., vol. 24, 1953, p. 344.

8. SUBHARMONIC OSCILLATIONS IN NONLINEAR SYSTEMS, C. Hayashi. *Ibid.*, p. 521.

9. A Mathematical Analysis of a Dielectric Amplifier, L. A. Pipes. *Journal of Applied Physics*, New York, N. Y., vol. 23, no. 8, Aug. 1952, pp. 818–24.

10. A Mathematical Analysis of a Series Circuit Containing a Nonlinear Capacitor, Louis A. Pipes. *AIEE Transactions*, vol. 72, pt. I, July 1953, pp. 238–44.

11. Applications of Integral Equations to the Solution of Nonlinear Electric Circuit Problems, Louis A. Pipes. *Ibid.*, Sept., pp. 445–50.

12. Graphical Analysis of Nonlinear Circuits Using Impedance Concepts, J. S. Thomsen. *Journal of Applied Physics*, New York, N. Y., vol. 24, 1953, p. 1379.

13. Bibliography of Magnetic Amplifier Devices and the Saturable Reactor Art, James G. Miles. *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 2104–23.

14. Comparison of Methods of Analysis of Magnetic Amplifiers, L. A. Finzi, G. F. Pitman, Jr. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 8, 1952, p. 144.

◆

No Discussion

The Cyclic Integrator—A Device for Measuring the Frequency Response of Magnetic Amplifiers

T. DUNNEGAN, JR.
ASSOCIATE MEMBER AIEE

J. D. HARNDEN, JR.
ASSOCIATE MEMBER AIEE

THE increasing number of applications of magnetic amplifiers to automatic control problems has brought about the necessity of being able to measure and to describe analytically the performance of such amplifiers as a function of time, so that the designer is able to predict the stability of a control system in which they are used. The time function most useful in applying a magnetic amplifier to a closed-loop control system is its frequency-response characteristic which describes the action of the amplifier in terms of amplification and phase shift of a sinusoidal control voltage which is varied in frequency over some specified range.

This article describes equipment and a technique that have been developed to measure magnetic amplifier frequency response. An analytical expression for the transfer function of a full-wave self-saturating magnetic amplifier is proposed on the basis of experimental data presented.

Cyclic Integration

Obtaining frequency-response data for a magnetic amplifier is complicated by its pulsating, variable wave form, output voltage. Interpreting such a voltage is difficult, since the wave form peak value is not a measure of its average value. The component of output voltage having control-signal frequency must be extracted from this wave form and the gain and phase shift between it and the control signal determined. The average value of output voltage is the quantity usually of prime interest because most electrical control elements are linearly responsive to their average input voltage.

An electronic measurement device, called a cyclic integrator has been developed which transforms the magnetic amplifier output voltage wave form to a different wave form that can be easily interpreted. The cyclic integrator generates a narrow pulse of voltage during each half-cycle of the power supply voltage. The magnitude of each pulse is proportional to the average value of the magnetic amplifier output voltage during the half-cycle period preceding the pulse. The envelope of this series of pulses can be observed with an oscilloscope to study the manner in which the average output voltage of the magnetic amplifier varies. The equipment and measurement method can be used with any magnetic amplifier circuit which produces an output voltage of the general wave shape shown as e_f in Fig. 1. That is, the output is a rectified sine wave in which only the leading part of each pulsation is varied to control the output.

Fig. 1 shows the basic circuit for the cyclic integrator. The integrating circuits are the resistance-capacitance networks R_1, C_1, and R_2, C_2. The time constants of these circuits are of equal value and are sufficiently long so that the voltage across the capacitors approximates very closely the time integral of the magnetic amplifier output voltage, for the proper relation between R and C as indicated by the approximation in Fig. 2. A twin-triode tube, properly phased, controls the discharge of capacitors C_1 and C_2 during each cycle of carrier voltage so that the average value of each output pulse results.

In connection with the wave forms shown in Fig. 1, at $t=t_2$

$$e_1=\frac{1}{R_1C_1}\int_{t_1}^{t_2} e_f dt \qquad (1)$$

At t_2 the gating voltage e_3 goes positive, driving the left-hand side of the twin triode to zero bias, and causing a rapid discharge of C_1 as shown by the wave form of e_1. At the same time the gating voltage e_5 goes negative allowing C_2 to charge.

The rapid change in e_1, impressed on the differentiating circuit R_3, C_3 which has a very short time constant, produces a pulse of voltage e_2 across R_3 equal in magnitude to the voltage e_1 at t_2. Thus

$$|e_2|=e_1=\frac{1}{R_1C_1}\int_{t_1}^{t_2} e_f dt \qquad (2)$$

The average value of any voltage as a function of time, between two limits t_1 and t_2 is

$$e_{av}=\frac{1}{t_2-t_1}\int_{t_1}^{t_2} e dt \qquad (3)$$

Therefore, the magnitude of the first pulse at t_2 is

$$|e_2|=\left(\frac{t_2-t_1}{R_1C_1}\right)e_{fav} \qquad (4)$$

Since t_2-t_1 is one half-cycle of the power supply frequency and is a constant

$$e_2=\text{Constant}\times(\text{average value of } e_f \text{ between } t_1 \text{ and } t_2) \qquad (5)$$

During time t_2 to t_3, capacitor C_2 charges, making

$$e_4=\frac{1}{R_2C_1}\int_{t_2}^{t_3} e_f dt \qquad (6)$$

Paper 54-260, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted February 8, 1954; made available for printing April 27, 1954.

T. Dunnegan, Jr., is with Chance Vought Aircraft, Inc., Dallas, Texas, and J. D. Harnden, Jr., is with the General Electric Company, Schenectady, N. Y.

The development and experimental work covered in this paper was performed at the General Engineering Laboratory of the General Electric Company, Schenectady, N. Y. The assistance of G. C. Dodson and L. R. Pangburn in the preparation of the paper is greatly appreciated.

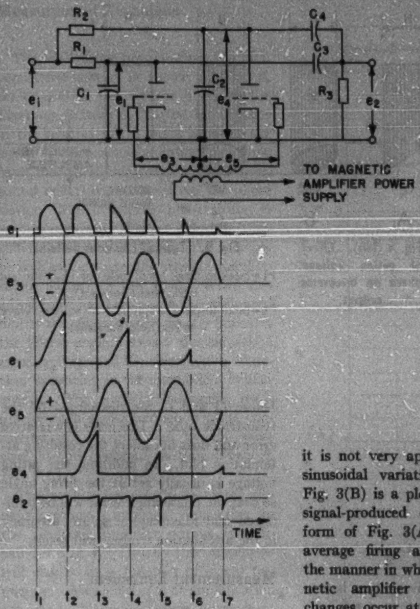


Fig. 1. Basic cyclic integrator circuit and wave forms

and at t_3, the gating voltage e_5 goes positive, driving the right-hand side of the twin triode to zero bias, and short-circuiting C_2. This rapid change in e_4 is impressed on the short time constant circuit R_3, C_4; and a pulse of voltage is produced across R_3 equal in magnitude to the voltage e_4 at t_3. By the same analysis used before, the magnitude of the second pulse is

$$e_2 = \text{Constant} \times (\text{average value of } e_1 \text{ between } t_2 \text{ and } t_3) \quad (7)$$

This process is repeated so that at the end of each half-cycle of the gating voltage, a pulse is produced which has a magnitude proportional to the average value of the input voltage during the preceding half-cycle. Thus a magnetic amplifier output wave form, applied as e_1 to the cyclic integrator, is transformed to a series of pulses which varies in the same manner as the magnetic amplifier average output voltage.

Fig. 3(A) shows the output voltage of a magnetic amplifier, the firing angle of which is being modulated either side of an arbitrary 90-degree firing angle by a sinusoidal signal. From the wave form it is not very apparent that there is a sinusoidal variation in its magnitude. Fig. 3(B) is a plot that shows only the signal-produced changes in the wave form of Fig. 3(A) measured from the average firing angle. It demonstrates the manner in which changes in the magnetic amplifier output occur. These changes occur at discrete time intervals and can be represented as small pulses of voltage either adding to or subtracting from some average output voltage These pulses have variable width, and the magnitude of their effect on the unmodulated wave form is difficult to visualize. However, if each variable width pulse is replaced by a pulse having a peak value proportional to the volt-time integral of the variable width pulse, a wave form such as that shown in Fig. 3(C) will result. The envelope of this wave form follows the wave shape of the input signal and shows the sinusoidal nature of the magnetic amplifier output variations.

The cyclic integrator output wave form, shown in Fig. 3(D), is composed of pulses similar to those represented in Fig. 3(C). The cyclic integrator pulses show the manner in which the incremental changes add to and subtract from the unmodulated magnetic amplifier output.

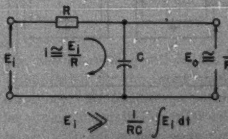


Fig. 2. Simple integrating circuit

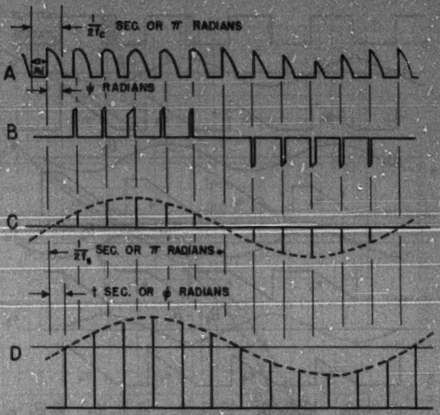


Fig. 3. A—Modulated magnetic amplifier output voltage. B—Pulses representing incremental change in wave form. C—Average value of incremental pulses. D—Cyclic integrator output

It must be noted that for the case shown in Fig. 3(D), the cyclic integrator pulses occur at the end of the power supply half-cycle. For this condition there is a relative phase shift between the ideal envelope of the pulses of Fig. 3(C) and the actual envelope of the cyclic integrator pulses of Fig. 3(D). This phase shift amounts to an error in the cyclic integrator indication. The time delay between the two envelopes is indicated on these figures by time t, which can be expressed as a phase shift in terms of either the power supply frequency or the signal (modulation) frequency.

In terms of the power supply frequency

$$t = \frac{\psi}{\pi} \frac{1}{2f_c} \quad (8)$$

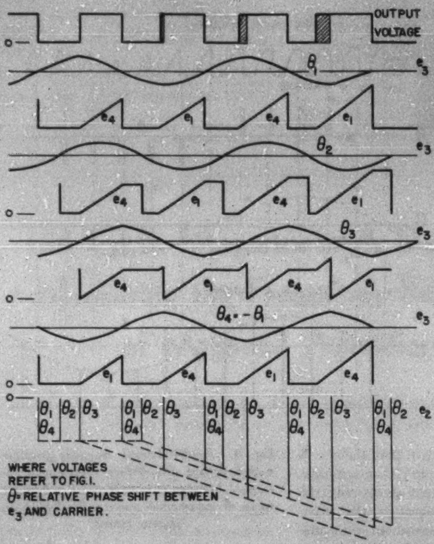
where ψ is the phase angle between the firing point of the magnetic amplifier and the occurrence of the cyclic integrator output pulse in radians and f_c is the power supply frequency in cycles per second.

In terms of the signal frequency

$$t = \frac{\phi}{\pi} \frac{1}{2f_s} \quad (9)$$

where ϕ is the phase angle between the envelope of incremental change and the envelope of cyclic integrator output and f_s is the modulation frequency in cycles per second. Thus

$$\frac{\phi}{\pi} \frac{1}{2f_s} = \frac{\psi}{\pi} \frac{1}{2f_c} \quad (10)$$



WHERE VOLTAGES REFER TO FIG. 1.
θ = RELATIVE PHASE SHIFT BETWEEN e_3 AND CARRIER.

Fig. 4 (left). Effect of gating voltage phase on integrator output

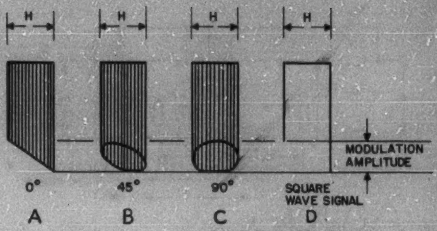


Fig. 6. Typical Lissajous patterns

H = peak-to-peak amplitude of $E_s \sin \omega_s t = 2E_s$ (held constant)

H = peak-to-peak amplitude of square wave $= 2E_s$ (held constant)

$$\phi = \frac{f_s}{f_c}\psi \qquad (11)$$

The phase shift error ϕ between the envelope of the cyclic integrator output and the actual modulation of the magnetic amplifier output is therefore proportional to the ratio of the signal frequency to the power supply frequency and to the phase angle ψ between the firing point of the magnetic amplifier and the occurrence of the cyclic integrator output pulses. This error is made negligibly small by reducing ϕ to a minimum through the use of a phase shift network in the gating voltage circuit.

In Fig. 4 a square wave output is considered to clarify the action of the integrator. The gating voltage e_3 is varied in phase so that voltage e_1 and e_4 across capacitors C_1 and C_4 vary. However, the composite curve of voltage e_2 shows that approximately the same peak value results for all cases, while the amount of phase error depends on the position of the gating voltage.

Some positions of gating voltage will result in more accurate integration than others, depending on the length of time before C_1 and C_2 are discharged by the twin-triode tube. This error is a relative error and does not affect the resulting attenuation plot. In addition, the gating voltage is usually set at the firing angle so that the phase-angle error will be minimized and maximum absolute accuracy in the attenuation reading will result.

Measurement Equipment

The block diagram of the system used for measuring the frequency response of the magnetic amplifier is shown in Fig. 5. An electronic ultralow-frequency oscillator is used to generate a signal voltage which is amplified by a direct coupled power amplifier. The output of the power amplifier drives a signal winding on the magnetic amplifier through an attenuator, resulting in an output which is modulated at the signal frequency. The output voltage from the magnetic amplifier is applied to the cyclic integrator and thereby transformed to a wave form of pulses, the amplitude of which varies in accordance with the average output voltage. The output of the cyclic integrator is applied to the vertical deflection amplifier of a d-c oscilloscope. The horizontal deflection amplifier of the oscilloscope is driven by the sinusoidal signal voltage from the oscillator after it has passed through a calibrated phase shifter. Driving the oscilloscope in this manner produces a Lissajous pattern which is used to indicate phase shift and amplification.

To achieve maximum safety in operation, one terminal of the equipment and magnetic amplifier output is grounded, as shown in Fig. 5. When the magnetic amplifier is supplied from a grounded a-c line, an isolation transformer is required to prevent short-circuiting part of the magnetic amplifier circuit.

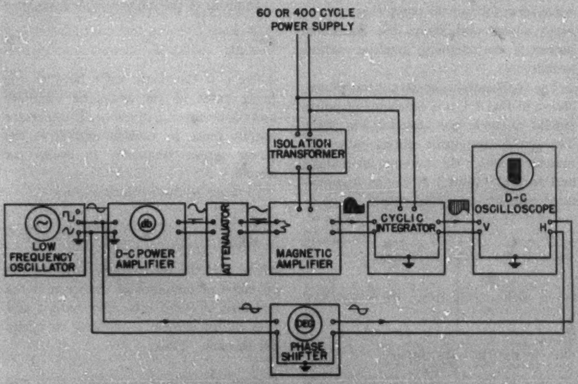


Fig. 5. Block diagram for cyclic integrator equipment

The oscilloscope presentation on a high persistence screen is shown in Fig. 6, demonstrating the manner in which the Lissajous pattern is effected by changes in the relative phase angle between the signal frequency modulation envelope from the cyclic integrator and the horizontal deflection voltage from the phase shifter. The adjustable phase shifter is used to reduce the relative phase shift between the vertical and horizontal voltages to zero as shown by the straight-line Lissajous pattern in Fig. 6(A). When this adjustment has been made the calibrated dial of the phase shifter directly indicates the phase angle between the sinusoidal signal voltage and the sinusoidal modulation appearing in the magnetic amplifier output. The correct phase-angle reading for each signal frequency value considered will be indicated by the same straight-line Lissajous pattern shown in Fig. 6(A).

The modulation amplitude of the signal-frequency component appearing in the magnetic amplifier output is indicated in Fig. 6. Usually this amplitude will change as a function of signal frequency. Measurements of amplification are made by maintaining a constant modulation amplitude each time the signal frequency is changed. This is accomplished by changing the driving power to the magnetic amplifier signal winding an appropriate amount by means of a gain control in the power amplifier. Before each reading this gain control, which is calibrated in decibels, must be adjusted to give a constant modulation amplitude for each signal frequency as indicated by the trapezoidal oscilloscope pattern. It is to be noted that the gain control affects only the a-c component and not the d-c bias level.

A value of amplification for steady-state conditions (i.e., for d-c signals), in addition to those values obtained for sinusoidal signals, is necessary to complete the frequency-response curve. This value may be obtained by using the square-wave output available from the low-frequency oscillator. A square wave with a peak-to-peak amplitude of $2E_s$ is used instead of the sine wave $E_s \sin \omega_s t$. The frequency of the square wave is made sufficiently low so that the magnetic amplifier output reaches a steady-state value after each reversal of the square wave. Then the calibrated gain control of the power amplifier is adjusted so that the same modulation amplitude that was used for the sine wave signal test is obtained using the square wave signal. The decibel reading will be the relative amplification for d-c input signals zero frequency. An oscilloscope pattern similar to that shown in Fig. 6(D) will be obtained with the square wave input and the modulation amplitude is measured as shown.

Thus, points on a frequency-response curve for a given magnetic amplifier can be obtained with the system shown in Fig. 5 by adjusting the equipment to give a trapezoidal pattern of constant

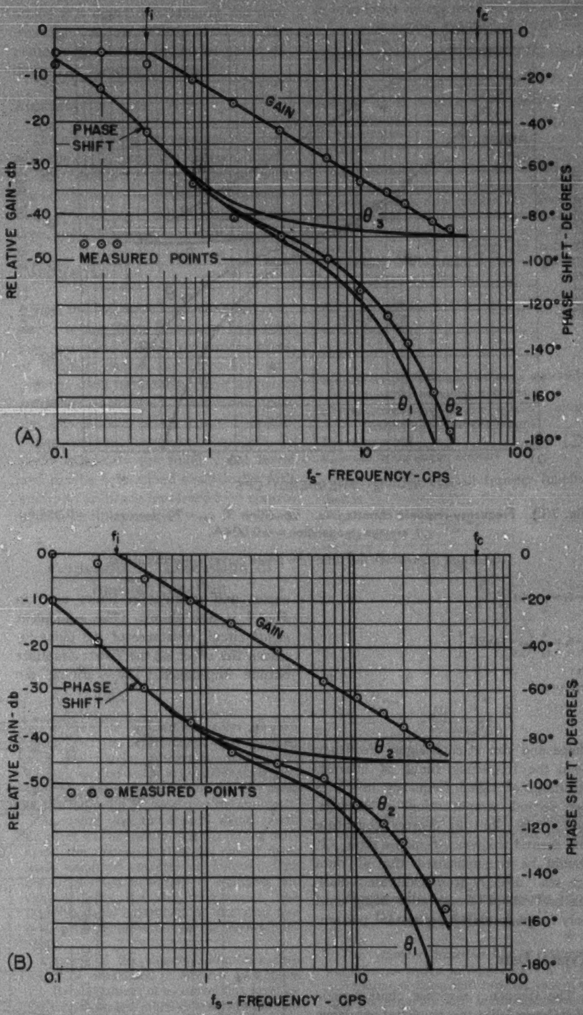


Fig. 7. Frequency-response characteristics. (A) Condition 1: $\alpha_{av}=141$ degrees; $I_L=0.015$A; average modulation $=\pm 0.004$A. (B) Condition 2: $\alpha_{av}=108$ degrees; $I_L=0.035$A; average modulation $=\pm 0.004$A. See Fig. 7(C) on next page for equations from which the solid-line phase curves were calculated

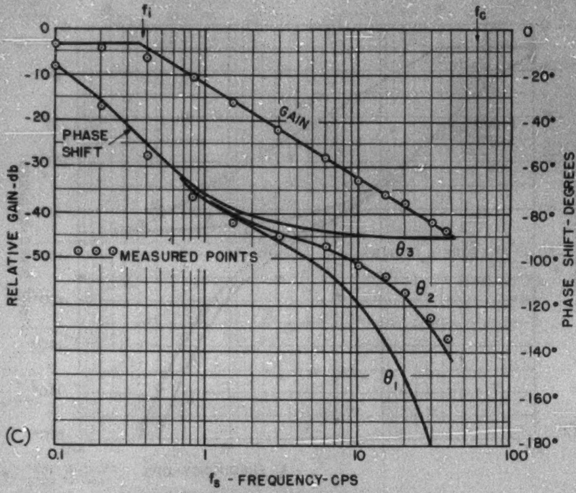


Fig. 7(C). Frequency-response characteristic. Condition 3: $\alpha_{av} = 78$ degrees; $I_L = 0.055A$, average modulation $= \pm 0.004A$.

Equations from which the solid-line phase curves were calculated:

$$-\theta_1 = \frac{f_s}{f_c}\pi + \tan^{-1}\frac{f_s}{f_t}$$

$$-\theta_2 = \frac{f_s}{f_c}\alpha_{av} + \tan^{-1}\frac{f_s}{f_t}$$

$$-\theta_3 = \tan^{-1}\frac{f_s}{f_t}$$

shape and size at each signal frequency over the frequency range of interest. Readings of signal frequency in cycles per second, amplification in decibels, and phase shift in degrees constitute the desired frequency-response data. It should be re-emphasized that all data are read directly from calibrated dials, the oscilloscope presentation being used only for maintaining a constant pattern.

Typical Data

The frequency-response characteristic of a General Electric amplistat catalogue no. *93G753G1*, is summarized in Fig. 7.

This is a one volt-ampere, plug-in, self-saturating magnetic amplifier using square loop core material and germanium rectifiers. The frequency response curves are for three different firing angles and small amounts of modulation.

The static transfer characteristic for the amplifier is shown in Fig. 8. This curve was obtained in the normal manner by varying the d-c signal only, with zero modulation. The three average firing angles and modulation limits used in Fig. 7 is also shown. The equivalent signal circuit resistance of 123 ohms includes the effect of the power amplifier output impedance. The variable d-c voltage source indicated has zero internal impedance and represents the bias adjustment which is incorporated in the power amplifier.

In Fig. 7 the N^2/R of the signal circuit was held constant while the firing angle was varied. The circles in these figures represent the experimental attenuation and phase-angle points as determined by the cyclic integrator equipment. The break frequency, f_t, for each of the three conditions was determined graphically by finding the intersection of the low-frequency and high-frequency asymptotes. In addition to the experimental data there are three phase curves shown which were calculated from equation 12, using the value of f_t as indicated in the figures.

$$\frac{E_0}{E_s} = \frac{K\epsilon^{-j\frac{\omega_s}{\omega_c}\alpha_{av}}}{\left(1+j\omega_s\frac{L_s}{R_s}\right)} = \frac{K\epsilon^{-j\frac{f_s}{f_c}\alpha_{av}}}{(1+j2\pi f_s T_t)} \qquad (12)$$

where

K = a constant

$T_t = \frac{1}{2\pi f_t}$

$f_s = \omega_s/2\pi$ = signal frequency
$f_c = \omega_c/2\pi$ = carrier frequency
α_{av} = the average firing angle, radians
E_0 = the output voltage change (a sinusoidal quantity of signal frequency)
E_s = the input voltage producing the change (a sinusoid)

The equation consists of a term in the

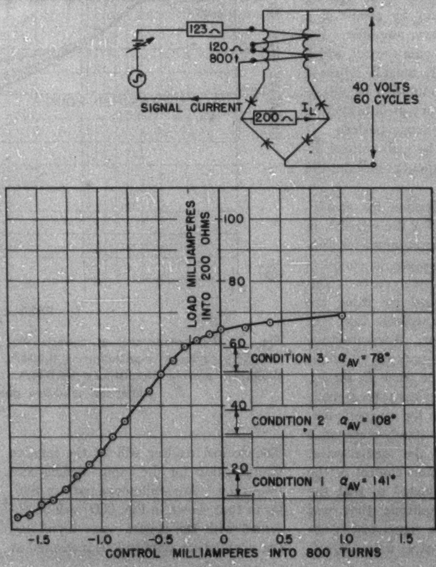


Fig. 8. Transfer characteristic of General Electric amplistat catalogue no. 93G753G1

denominator which is associated with the normal inductance-resistance signal circuit time constant[1] and a term in the numerator which is called the transportation lag. The phase angle associated with the first term is θ_3, while θ_1 and θ_2 are the total phase shift of the transfer function calculated for two conditions of the transportation lag. The assumption is often made that this lag is a full half-cycle of carrier frequency indicating that the firing angle is 180 degrees or π radians. This applies for the θ_1 curve in which α has been set equal to π radians and indicates the maximum phase shift. However, the best correlation between calculated and experimental phase-angle curves results when the transportation lag phase shift is a function of the firing angle α as for the θ_2 curve. The agreement of the analytical and experimental curves indicates the usefulness of the transfer function which is nonrigorously derived in Appendixes I and II.

Additional information regarding the time constant of the amplistat can be obtained from the data shown in Fig. 7. At the three points checked experimentally the time constant, as determined from the break frequency f_t, is longer for $\alpha=108$ degrees than it is for $\alpha=78$ or 141 degrees. This indicates that the time constant varies with the firing angle α being maximum in the linear part ($\alpha=\pi/2$ radians) of the control characteristic. This might be thought of as the point of maximum effective inductance. It is of particular interest to note the correlation between the conclusion of Storm[2] and these experimental data.

Versatility of Cyclic Integrator

A complete analysis of the magnetic amplifier circuit for widely varying conditions is readily obtained with the equipment described in the paper because the accumulation of test data is greatly facilitated. This is largely a result of:

1. Direct reading attenuation and phase angle dials.
2. Rapid adjustment of equipment at each new test frequency.
3. Accuracy of measurements.

As an example, the common time constant versus N^2/R curves[2] is easily plotted from a series of frequency response curves. All of the curves would be taken at the same firing angle and with the same number of signal turns, while the signal circuit resistance would be changed in each curve. In each resulting frequency-response curve the signal frequency at which the break in attenuation curve occurs is a direct indication of the time constant. As an additional check the phase-angle reading should be about 45 degrees at the same frequency.

Accuracy

The accuracy of the readings obtained with the cyclic integrator depends largely on the familiarity of the operator with the equipment. In general, his measurement of phase angle will be within ±3 degrees and attenuation within ± 1 decibel.

Conclusions

The use of extremely nonlinear elements in the form of square hysteresis loop core materials and the various types of rectifiers makes the servo synthesis problem seem formidable where magnetic amplifiers are involved. Indications from the experimental study covered in this paper are that magnetic amplifiers do perform linearly for small signal levels and can be determined analytically with sufficient accuracy for closed-loop control system stability studies.

Appendix I. Discussion of Transfer Function

The expression

$$\frac{E_0}{E_s}=\frac{K\epsilon^{-j\frac{\omega_s}{\omega_c}\alpha_{av}}}{\left(1+j\omega_s\frac{L_s}{R_s}\right)}=\frac{K\epsilon^{-j\frac{f_s}{f_c}\alpha_{av}}}{(1+j2\pi f_s T_1)} \tag{12}$$

has not been derived but does correlate very closely with the experimental data obtained with the cyclic integrator equipment. Though no derivation is available, it is possible to deduce the expression on the basis of the following physical picture approach.

Assuming that the control signal exerts no effect on the core during the gating half-cycle, the conventional bridge circuit magnetic amplifier operates as follows: The output voltage is varied by changing the time of firing, (i.e. the firing angle); the firing angle is controlled by the core flux level prior to initiation of the gating half-cycle.

Control of the magnetic amplifier then might be considered to occur in two steps: 1. Establishment of a control flux level by the input voltage, much as occurs in the conventional iron core reactor. 2. Expression of this control level in the output after a time delay imposed by the magnitude of the firing angle.

The control flux level, being akin to flux change in a conventional reactor might then be related to the input voltage by

$$\frac{\Delta\phi}{E_s}=\frac{K_1}{\left(1+j\omega_s\frac{L_s}{R_s}\right)}=\frac{K_1}{(1+j\omega_s T)} \tag{13}$$

where

$\Delta\phi$ = the change in control flux level
E_s = the input voltage producing the change (a sinusoidal quantity of signal frequency)
L_s = the equivalent inductance of the signal circuit
R_s = the signal circuit resistance
ω_s = the signal frequency in radians per second
K_1 = a constant
$T=L_s/R_s$

Expression of the control flux in the output, is delayed by the time required for gating voltage (volt-time integral) to effect the flux excursion from control flux level to saturation. This might be considered a transportation lag. The relationship between control flux level and expression of the control in the output then is

$$\frac{E_0}{\Delta\phi}=K_2\epsilon^{-j\frac{\omega_s}{\omega_c}\alpha_{av}} \tag{14}$$

where

E_0 = the output voltage change (a sinusoidal quantity of signal frequency)
ω_c = the frequency of the carrier voltage (supply), radians per second
K_2 = a constant
α_{av} = the firing angle, radians

The resulting over-all transfer function then is

$$\frac{E_0}{\Delta\phi}\frac{\Delta\phi}{E_s}=\frac{E_0}{E_s}=\frac{K_2\epsilon^{-j\frac{\omega_s}{\omega_c}\alpha_{av}}}{K_1\left(1+j\omega_s\frac{L_s}{R_s}\right)}$$

$$=\frac{K\epsilon^{-j\frac{\omega_s}{\omega_c}\alpha_{av}}}{\left(1+j\omega_s\frac{L_s}{R_s}\right)}$$

or (15)

$$\frac{K\epsilon^{-j\frac{f_s}{f_c}\alpha_{av}}}{\left(1+j2\pi f_s\frac{L_s}{R_s}\right)}=\frac{K\epsilon^{-j\frac{f_s}{f_c}\alpha_{av}}}{(1+j2\pi f_s T)}$$

where

K = a constant
f_s = the signal frequency, cycles per second
f_c = the carrier frequency, cycles per second
$T=L_s/R_s$

Appendix II. Expression for "Transportation Lag"

The transfer function for a magnetic amplifier as stated in Appendix I is of the nonminimum phase shift type because of the transportation lag term. An expression for this term follows in which the modulation is considered to be extremely small so that the change in firing angle is negligible.

In connection with Fig. 3:

α_{av} = firing angle in radians and can be made any value from 0 to π of the carrier
= time equivalent of the average firing angle and is the transportation lag in seconds

Also

$$\frac{\frac{1}{2f_c}}{\pi}=\frac{\tau}{\alpha_{av}} \tag{16}$$

or

$$\tau = \frac{\alpha_{av}}{2\pi f_c} = \frac{\alpha_{av}}{\omega_c} \tag{17}$$

Generally stated, the output of the magnetic amplifier is a function of time t and is equal to some function of the input which is a function of $(t-\tau)$ or

$$f_0(t) = f_i(t-\tau) \tag{18}$$

A resolution of this equation can be made by putting both functions on the same time base. To obtain f_i, t then express it in Taylor's series form

$$f_i(t) = f_i(a) + \frac{pf_i(a)(t-a)}{1!} + \frac{p^2f_i(a)(t-a)^2}{2!} + \ldots \tag{19}$$

where p is the differential operator.

Now let

$$a = (t-\tau) \tag{20}$$

Substituting

$$f_i(t) = f_i(t-\tau) + \frac{pf_i(t-\tau)\tau}{1!} + \frac{pf_i(t-\tau)\tau^2}{2!} + \ldots = f_i(t-\tau)\left[1 + \frac{p\tau}{1!} + \frac{p^2\tau^2}{2!} + \ldots\right] \tag{21}$$

$$f_i(t) = f_i(t-\tau)\epsilon^{p\tau}$$

Substituting in equation 18

$$f_0(t) = f_i(t)\epsilon^{-p\tau} \tag{22}$$

or the transfer function then becomes

$$\frac{f_0(t)}{f_i(t)} = \epsilon^{-p\tau} \tag{23}$$

This expression in terms of the measured quantities is

$$\frac{f_0(t)}{f_i(t)} = \frac{E_0}{\Delta\phi} = K_2\epsilon^{-p\tau} \tag{24}$$

But for a differential equation of this type with a sinusoidal driving function it can be written

$$\frac{E_0}{\Delta\phi} = K_2\epsilon^{-j\omega_s\tau} \tag{25}$$

Substituting equation 17 yields the expression for transportation lag

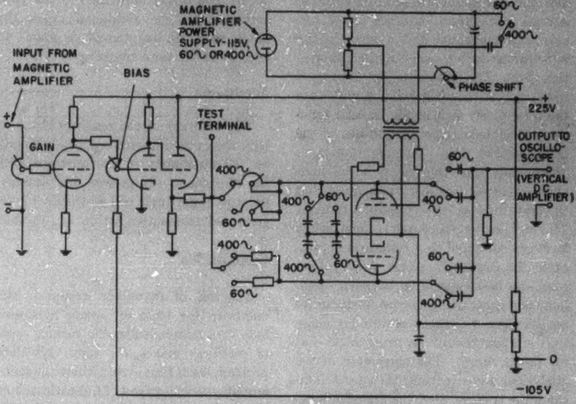
$$\frac{E_0}{\Delta\phi} = K_2\epsilon^{-j\frac{\omega_s}{\omega_c}\alpha_{av}} \tag{14}$$

Appendix III

Cyclic Integrator Equipment Diagram

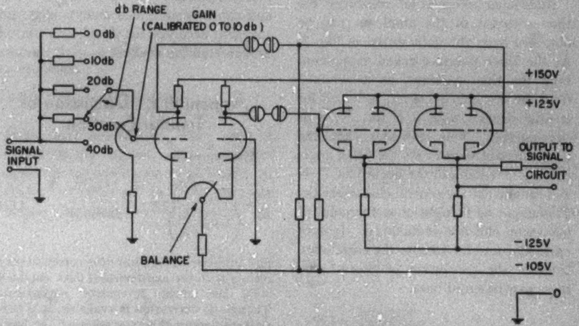
The equipment was designed to be used with 400-cycle and 60-cycle magnetic amplifiers. However, by changing simple time constant circuits, other frequencies could readily be used.

The power supplies used with the equipment must be well regulated and filtered if



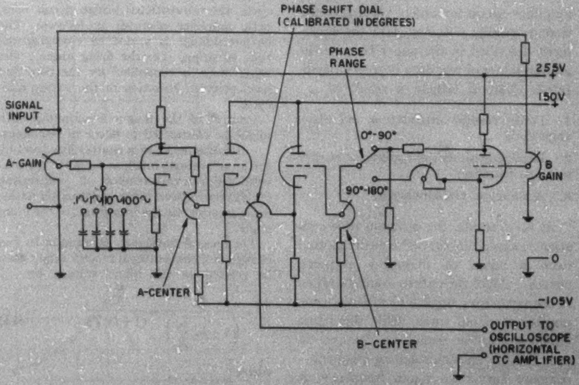
MAGNETIC AMPLIFIER POWER SUPPLY-115V, 60∿ OR 400∿
INPUT FROM MAGNETIC AMPLIFIER
BIAS
GAIN
TEST TERMINAL
60∿
400∿
PHASE SHIFT
+225V
OUTPUT TO OSCILLOSCOPE (VERTICAL D C AMPLIFIER)
400∿
60∿
0
-105V
+
-

Fig. 9. Cyclic integrator circuit diagram



db RANGE
GAIN (CALIBRATED 0 TO 10 db)
0db
10db
20db
30db
40db
SIGNAL INPUT
BALANCE
+150V
+125V
OUTPUT TO SIGNAL CIRCUIT
-125V
-105V
0

Fig. 10. Power amplifier circuit diagram



PHASE SHIFT DIAL (CALIBRATED IN DEGREES)
PHASE RANGE
255V
150V
SIGNAL INPUT
A-GAIN
0°-90°
90°-180°
B GAIN
400∿
A-CENTER
B-CENTER
0
-105V
OUTPUT TO OSCILLOSCOPE (HORIZONTAL D C AMPLIFIER)

Fig. 11. Phase shifter circuit diagram

maximum performance from the device is to be realized.

Cyclic Integrator

The diagram for the cyclic integrator shown in Fig. 9 is built around the basic circuit considered in Fig. 1. A phase shift network is added between the magnetic amplifier power supply voltage and the cyclic integrator tube grid transformer, providing a means of shifting the integrator output pulses relative to the magnetic amplifier firing angle.

Power Amplifier

Fig. 10 is of interest because it shows the nature of the driving impedance seen by the magnetic amplifier and the nature of the attenuation networks. The balance control varies the d-c level of the power amplifier output voltage, providing a convenient bias source for the signal winding.

Phase Shifter

An examination of Fig. 11 indicates that the phase shifter depends on the generation of two voltages *A* and *B*, at input frequency, and displaced in time by 90 electrical degrees. In operation, the magnitudes of voltages *A* and *B* are made equal so that by connecting a potentiometer between them, a voltage of any intermediate phase is available. Therefore, a voltage of input frequency but variable in phase is available at the output terminals.

References

1. The Figure of Merit of Magnetic Amplifiers, J. T. Carleton, W. F. Horton. *AIEE Transactions*, vol. 71, pt. I, Sept. 1952, pp. 239–45.

2. Theory of Magnetic Amplifiers with Square-Loop Materials, H. F. Storm. *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 629–40.

3. Progress Report of the AIEE Magnetic Amplifier Subcommittee. AIEE Committee Report. *AIEE Transactions*, vol. 70, pt. I, 1951, pp. 445–50.

4. The Application of a 60-Cycle Magnetic Amplifier to a Position Indicating Servomechanism, Albert E. Schmid. *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 7, 1951, pp. 515–22.

◆

No Discussion

Precision Measurement of Phasors in Electric Networks as a Function of Frequency

A. J. HERMONT
MEMBER AIEE

Synopsis: Measurement of phasor relationships in networks is considered in general, and a system is described which accomplishes this task. Numerical examples for purposes of orientation are given. Detection is based on the performance of a phase-sensitive rectifier of the ring modulator type. A system is described, using a pair of such detectors being polarized by two auxiliary voltages, of the frequency under consideration, the voltages being in time quadrature to each other. In this manner true null detection is accomplished unambiguously. The differential polarized rectifier, the polarizing supply circuit, and the required auxiliary amplifiers, containing means of amplitude-and-phase adjustment, are discussed briefly. Operation in the frequency range of 1 to 1,000 cycles for input signals of the order of 100 microvolts and for extremely adverse signal-to-noise ratios is accomplished with an over-all accuracy of 1 per cent.

THE problem of precision measurement of phasor relations in circuits of the general type, both passive and active, 2-pole and 4-pole, depends in the last analysis on the accurate determination of magnitude and phase. Magnitude may now be measured with any desired accuracy since calibrated attenuators are available with errors of about 1 per cent up to 50 megacycles,[1] while difficulties in the audio range are easily overcome by adequate design. It is the precision measurement of phase which has caused delay in a broad field of activity. It is surprising that progress in this direction has not been made until relatively recently.

The process of measuring electrical phasors, quite naturally, may be subdivided into two parts: the part which deals with the system accomplishing the necessary (calibrated) magnitude and phase changes, and the part which deals with actual detection of the desired condition. When both parts are grouped into one, previous work in this field may be briefly summarized as follows.

There are at present many principles on which phase measurement is based. The "sum-and-difference" system[2–5] is based on the fact that the vector sum of two sinusoidal vectors of equal amplitude but differing in phase by an angle φ is proportional to cos $\varphi/2$ while the difference is proportional to sin $\varphi/2$. Thus by using amplifiers with good automatic volume control and performing summation and subtraction, meter readings obtain which are trigonometric functions of the phase angle. A study of the sum and difference of rectified sine waves[6] has also been made and applied to the determination of small phase shifts.

The cathode-ray tube as a null indicator has been used for quite some time in the Lissajous mode. If the output of a network is connected to the vertical deflection plates and the input via a phase shifter to the horizontal plates, and if the phase shifter is adjusted to become an exact quadrature to, say, the imaginary component of the output voltage, then observation of the ellipse reveals information as to the adjustment of the imaginary component only, and balancing becomes unique.[7] The collapse of the ellipse to a straight line may be used to observe approach to zero phase shift.[8] More novel systems using the cathode-ray oscillograph are the following: A circular trace is used in conjunction with intensity modulation[9] such that the locus of maximum intensity depends on the phase of the modulating voltage: as a result a portion of the circle is dim, which is indicative of phase. Intensity control in the form of pulses[10] can also be used, in which case both amplitude and phase are displayed on the screen. In another scheme,[11] which works best at radio frequencies, d-c components are developed which are functions of cos φ and sin φ, and these, applied to the vertical and horizontal plates of the cathode-ray oscillograph, position the spot. If the spot is changed into a trace, this radial trace indicates phase directly.

Square-wave technique provides another means to measure phase and amplitude; because of the inherent high-frequency content of the wave, the transient response is analyzed for the desired information.[12] In a steady-state manner,

---

Paper 54-292, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted October 23, 1953; made available for printing April 27, 1954.

A. J. Hermont is with the Shell Development Company, Houston, Tex.

Gratitude is expressed to the Shell Development Company for permission to publish this paper.

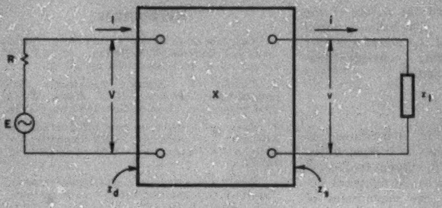


Fig. 1 (above). Definition of quantities

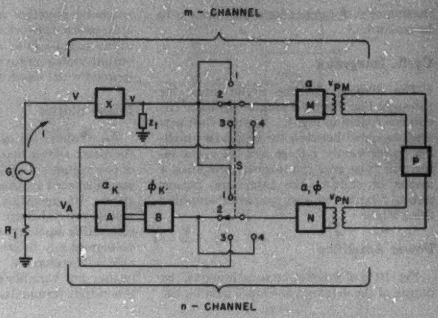


Fig. 2. Measurement of transfer impedance

employing pulse technique, the system transforms the two sine waves into square waves; the latter are differentiated, producing pips which trigger a flip-flop circuit. In the end, average current flows which is proportional to the fraction of the time between pips and thus to the phase difference between the two sinusoidal waves.[12,14] In a modification of this idea biased limiters are used to produce trapezoidal waves.[15]

Another principle utilizes a rectification scheme, employing two diodes and transformer coupling, in which there is developed a d-c component which is proportional to cos φ (cosine galvanometer).[11,16,17] In a modification of this, triodes may also be used.[18] The ring modulator, acting as another form of the cosine galvanometer, may be used as the phase-sensitive device.[7,19] A somewhat different aspect of this principle is also possible.[20] Measurement is greatly facilitated if it is done at a single frequency.[21] This idea is also the basis of the system referred to previously.[2]

A special 9-electrode tube, the φ-detector,[22] should be mentioned: plate current can flow only when the potential is positive at two of the grids simultaneously, so that the current is an almost linear function of the phase between the two grid potentials.

All these systems have their advantages and disadvantages. The development took two main directions: the desirability to create systems which could produce rapid, direct reading information, for a restricted frequency band, under sacrifice of precision,[4,5,9–11,13,15] and the desirability to have available laboratory-type instruments with great inherent accuracy, such as the one developed by the Bell Laboratories.[2] The system suggested by Columbia University[6] is far less complex and yet sufficiently accurate for the frequency range 20 cycles to 100 kc.

The measuring system presented in the following is based on the ring modulator as the phase-sensitive device. It was specifically developed for the precision measurement of transfer impedance in the frequency range 1 to 1,000 cycles under adverse conditions, i.e., signals in the order of 100×10^{-6} volt and extremely low signal-to-noise ratios. The accuracy of determination had to be within 1 per cent or better for the entire frequency range and for the entire 2π phase space. The detecting technique is such that the phasor is measured not in two steps involving amplitude and phase but rather as an invariant quantity in an auxiliary co-ordinate system.

The system does not lend itself easily to direct-reading interpretation, nor is the measuring process a rapid one. Stress is laid on accuracy, which is entirely determined by the accuracy of a few resistive and capacitive components, and on reliability. Design principles rather than design details are considered in the following discussion.

Basic Measurement Considerations

Consider the general situation illustrated in Fig. 1, where a generator of internal resistance *R* and electromotive force *E* feeds a 4-pole *X* which is terminated in an impedance z_t. The network *X* may be active or passive and may contain under certain conditions nonlinear elements. At the input terminals the electrical conditions are characterized by the phasors I, V, while the conditions at the output terminals are identified by the quantities i, v. The following phasor quantities associated with this configuration are specifically defined:

$z_d=V/I$ = driving point impedance, which conforms to the usual designation if *R* is considered as part of the generator.

$z_s=v'/i'$ = source impedance, which represents the equivalent output impedance of *X* with z_t disconnected.

$z_t=v/i$ = terminal impedance.

$z_I=v/I$ = transfer impedance, i.e., the quantity which transfers an input current to an output voltage.

$y_V=i/V$ = transfer admittance, i.e., the quantity which transfers an input voltage to an output current.

$\varrho_V=v/V$ = transmission ratio, i.e., the quantity which transfers an input voltage to an output voltage.

Depending on the nature of the investigation, it may become important to measure with great precision either or all of the six quantities just enumerated. It will be shown, however, that only three of these are independent and therefore need to be measured, namely z_I, y_V, and ϱ_V.

From the definition it follows that

$$z_d=\frac{V}{I}=\frac{z_I}{\varrho_V} \tag{1}$$

and

$$z_t=\frac{v}{i}=\frac{\varrho_V}{y_V} \tag{2}$$

It is understood of course that if z_t is completely known ϱ_V becomes a dependent quantity, namely $\varrho_V=y_Vz_t$.

There remains only the source impedance. Let a quantity measured with $z_t=\infty$ be affixed with a subscript *oc* (open circuit) and if measured under condition $z_t=0$, let the subscript be *sc* (short circuit). It is known that

$$i_{sc}=\frac{v}{z_s} \text{ and } v_{oc}=v$$

so that

$$z_s=\frac{v_{oc}}{i_{sc}} \tag{3}$$

If therefore z_I is measured for $z_t=\infty$ obtaining

$$z_{I_{oc}}=\frac{v_{oc}}{I_{oc}}$$

and y_V for $z_t=0$, i.e.

$$y_{V_{sc}}=\frac{i_{sc}}{V_{sc}}$$

it follows from equation 3 that

$$z_t=\frac{z_{I_{oc}}}{y_{V_{sc}}}\frac{I_{oc}}{V_{sc}}$$

Assuming further that care is taken to keep $V_{oc}=V_{sc}=V=$constant during the measurement, it follows that

$$\frac{I_{oc}}{V_{sc}}=\frac{I_{oc}}{V}=\frac{1}{z_{d_{oc}}}=\frac{\varrho_{V_{oc}}}{z_{I_{oc}}}$$

Thus

$$z_t=\frac{\varrho_{V_{oc}}}{y_{V_{sc}}} \tag{4}$$

While this method is reserved for cases in which z_t represents a circuit with complete isolation of the input from the output (such as provided by vacuum tubes) it is not suitable for passive networks: for the condition $V_{oc}=V_{sc}=V=$constant is not relizable because of the fact that, in general, the phase displacement of the V-phasor does not by itself remain constant, and erroneous results would obtain. For such networks it is best to reverse the flow of energy and, removing z_t, to replace it by the generator, then proceeding as if the driving point impedance were measured: it represents the source impedance.

Returning now to the primary quantities z_I, y_V, and ϱ_V, the principle of the measuring system will be discussed. Considering the block diagram, Fig. 2, and the measurement of z_I, the existing situation develops as follows.

An ungrounded variable-frequency generator G feeds the series combination of the network under test X and a suitable resistor R_I in parallel with which lies an auxiliary circuit A. It is seen that X and A are located in different "channels," termed the m- and n-channels respectively. The detector, the nature of which is discussed in detail elsewhere,[23] is the device P, and it will suffice at present to state that the indicator is a d-c galvanometer and that P is essentially a differential device, i.e., one measuring the difference between two vectors. M and N are essentially amplifiers, B is a phase shifter, and A is a gain-loss circuit. The 4-position ganged switch S is manipulated in the sequence 1-2-3-4.

The basic philosophy leading to the measuring technique will now be presented qualitatively in order to afford a clear understanding of what is being measured and why. Reference to Fig. 2 and examination of the situation for the selector switch in the various positions leads to the observations in the following paragraphs.

SELECTOR SWITCH IN POSITION 1

M and N are both fed by one and the same source, namely the output of X. Regardless of what the magnitude or phase of v is, P surely must indicate true null under these circumstances if the M-P-N system as a whole is considered, for here two identical voltage vectors are impressed at the inputs of M and N. If the response of P is different from zero (and there exists no reason why it generally should be zero), circuits M and N may nevertheless be adjusted to force the response of P to become zero. Accordingly the magnitudes of the M and N amplification and the phase of the N amplification are adjusted. This then is the first step of the measurement. It amounts essentially to a balance of the N-P circuit against that of the M-P circuit when the input to both is v. With this auxiliary balance guaranteed, the measurement proceeds as follows.

SELECTOR SWITCH IN POSITION 2

M still being fed by v, N is being fed by the voltage drop produced by the input current I flowing through a known resistor R_I, i.e., V_A, amplified by A and phase-shifted by B. Now, as soon as A and B are so adjusted that P again indicates null, it must be concluded that, for the second time, the inputs of M and N are being energized by identical voltages. This step therefore essentially balances v against $I\times R_I\times$unknown complex gain of the $A-B$ aggregate. If the complex gain could be read off absolutely, the measurement could be completed in two steps, for $z_I=v/I$ could be obtained by computation. Unfortunately, however, a transformer must be used in the $A-B$ aggregate, if it is remembered that the full 2π phase space may generally have to come into play, which makes the complex gain a function of frequency and rather inaccessible to direct computation. It was therefore deemed advisable to resort to two additional steps.

SELECTOR SWITCH IN POSITION 3

M and N are again being fed by one and the same source, namely V_A. It is again demanded that the output of P indicate true null. This (third) balance is achieved by readjusting the M and N circuits so that this step essentially balances the $N-P$ circuit against the $M-P$ circuit when the voltage to both is V_A.

SELECTOR SWITCH IN POSITION 4

M still being fed by V_A, N is being fed by V_A amplified by A and phase-shifted by B. When A and B are readjusted so that P once more indicates null, it must be concluded that, for the fourth time, the inputs to M and N are energized by identical voltages. This step essentially changes the unknown complex gain of A and B in cascade, as it existed at the end of step 2, into a known real gain of unity for the same aggregate.

It will next be shown by means of a quantitative study that it is thus possible to obtain z_I without knowledge of either the absolute gain of the $A-B$ aggregate or the absolute phase shift produced by the latter. This will be done by merely utilizing a calibrated gain-or-loss factor and a computable phase difference.

In Figs. 2 and 4 the symbol α_k indicates that means are provided to achieve "known" attenuation; φ_k indicates that phase shift is obtainable from the particular block; α, φ without the subscripts refer to the provision of unknown attenuation and phase shift.

Assuming X to be an amplifying device, and having selected the required frequency and an appropriate input level for X and for A, which implies having decided on R_I, switch S being in position 1, the M and N amplifiers are fed by one and the same voltage v, and equation 5 may be set up

$$v\tau_{M1}+v\underline{\tau_{N1}}=\delta_1 \tag{5}$$

The quantities τ are transfer factors; they include gain, loss, rectification, and phase sensitivity, and they transfer an input voltage phasor to a d-c galvanometer deflection. The gains of M and N are first set to a convenient value. M and N are next balanced against each other, i.e., the process $\delta_1\rightarrow0$ is accomplished by adjusting both the amplitude and the phase of N. This is indicated in equation 5 by underlining τ_{N1} as being the quantity which is operated on to achieve $\delta_1\rightarrow0$. When $\delta_1\equiv0$ it may be stated that the galvanometer deflection caused by the M amplifier just cancels that caused by the N amplifier, a situation which exists in all null methods. What has therefore been achieved is expressed by the relation

$$\tau_{M1}=-\tau_{N1} \tag{6}$$

Switch S is now thrown in position 2 and the complete m and n channels are in operation. Since by definition $v=Iz_I$, the output from the m channel would be $Iz_I\tau_{M1}$. The input to the A amplifier is IR_I. The vector gain or loss of A may be represented by $\mu_{A2}e^{j\theta_{A2}}$, i.e., by a magnitude μ_{A2} and a phase displacement θ_{A2}. The phase-shift network may, as will be shown later, have loss besides phase displacement, and transmission through it may therefore be expressed by $\lambda_{B2}e^{j\theta_{B2}}$, i.e., by a loss factor and a phase displacement. Thus

$$Iz_I\tau_{M1}+IR_I\underline{\mu_{A2}}e^{j\theta_{A2}}\lambda_{B2}\underline{e^{j\theta_{B2}}}\tau_{N1}=\delta_2 \tag{7}$$

The approach of $\delta_3 \to 0$ shall be accomplished by operating on the magnitude of A and the phase of B. At balance, and utilizing equation 6, one obtains

$$z_I = R_I \mu_{A2} \lambda_{B2} e^{j\theta_{A2}} e^{j\theta_{B2}} \tag{8}$$

This is not yet the desired form, since the magnitude and phase relations of A must be known to evaluate z_I, a requirement which greatly adds to the burden of measurement, and which, moreover, is not necessary if another pair of steps is introduced. Consider position 3 of switch *S*: *M* and *N* amplifiers are fed by one and the same voltage V_A, which may be considerably smaller than v, so that the gains of *M* and *N* may have to be increased accordingly

$$V_A \tau_{M3} + V_A \underline{\tau_{N3}} = \delta_3 \tag{9}$$

and $\delta_3 \to 0$ is established (after a preliminary choice of the magnitudes of τ_{M3} and τ_{N3}) by manipulation of the gain-and-phase controls of *N*. For $\delta_3 = 0$

$$\tau_{M3} = -\tau_{N3} \tag{10}$$

In position 4 of switch *S*, *M* is still energized by V_A, while *N* is energized by the output of *B*. Thus

$$V_A \tau_{M3} + V_A \underline{\mu_{A4}} e^{j\theta_{A4}} \lambda_{B4} \underline{e^{j\theta_{B4}}} \tau_{N4} = \delta_4 \tag{11}$$

The balancing process shall now be accomplished in exactly the same manner as in connection with equation 7. There is no reason to assume that the phase of A varies with magnitude, and in fact this can be avoided by proper design; hence $\theta_{A4} = \theta_{A2}$. As for the loss associated with B, a variation of it with phase is unavoidable (for certain instances) in the present design; however the loss of B is a known function of the phase of B. The magnitude μ_{A4} of A can be interpreted as the magnitude μ_{A2} times a calibrated gain or loss factor ζ_{Ak}, thus $\mu_{A4} = \mu_{A2}\ \zeta_{Ak}$. Hence, for $\delta_4 = 0$

$$1 = \mu_{A2} \zeta_{Ak} \lambda_{B4} e^{j\theta_{A2}} e^{j\theta_{B4}} \tag{12}$$

Combination of this with equation 8 results in

$$z_I = R_I \frac{1}{\zeta_{Ak}} \left(\frac{\lambda_{B2}}{\lambda_{B4}} \right) e^{j(\theta_{B2} - \theta_{B4})} \tag{13}$$

Letting

$$\varphi_{Bk} = \theta_{B2} - \theta_{B4}$$

$$\beta_{Bk} = \frac{\lambda_{B2}}{\lambda_{B4}} \tag{14}$$

one finally obtains

$$z_I = R_I \frac{\beta_{Bk}}{\zeta_{Ak}} e^{j\varphi_{Bk}} \tag{15}$$

and the quantity z_I is uniquely determined by the known resistor R_I and by known networks A and B. This is an important relation which shall be illustrated qualitatively by means of the vector diagram of Fig. 3. To simplify the work it is sufficient to study the situation up to the outputs of *M* and *N*, i.e., up to the voltages v_{PM} and v_{PN} (see Fig. 2) and to neglect the effect of *P* altogether. Thus instead of equations 6 and 10 one should think of $\gamma_{M1} = \gamma_{N1}$ and $\gamma_{M3} = \gamma_{N3}$, i.e., of that condition for which the vector gains of *M* and *N* have been brought to equality. Letting $\gamma_{M1} = \mu_{M1} e^{j\theta_{M1}}$ and $\gamma_{M3} = \mu_{M3} e^{j\theta_{M3}}$, where (for the sake of generality as well as for practical reasons) it has been deliberately assumed that $\theta_{M1} \neq \theta_{M3}$, the vector diagram, can be reconstructed by the aid of the identities indicated in Fig. 3. It should be noticed especially that $\theta_{A2} \equiv \theta_{A4}$ must be guaranteed by design of A (a task which may be easily achieved in practice). This phase angle has been assumed lagging or negative, the clockwise angles being leading or positive. In the first pair of operations it is obviously the goal to make vector 08 coincide with vector 03: this is done by adjusting the A magnitude from an arbitrary μ_{A2}' to μ_{A2} and by shifting the known phase from an arbitrary θ_{B2}' to a total shift $\theta_{B2} = \theta_{B2}' + \theta_{B2}''$. In the second pair of operations the result is achieved if vector 015 is made to coincide with 010. Here again the magnitude is adjusted from the arbitrary μ_{A4}' to μ_{A4}. The phase at the start of the second pair of operations is usually not arbitrary; rather it is the total B phase at the end of the first pair of operations, thus $\theta_{B4}' = \theta_{B2}$. It can be checked from the diagram that this makes the output of the B network 012 precisely in phase with the output voltage of network under test, $v(02)$

$$-\theta_{A2} + \theta_{B4}' = \Psi$$

The vector diagram will be used to prove the truth of equation 15 in this manner. Forming $\varphi_{Bk} = \theta_{B2} - \theta_{B4} = \theta_{B2} - (\theta_{B2} - \theta_{B4}'') = +\theta_{B4}''$ one finds by inspection that $\psi + \theta_{M3} = \theta_{B4}'' + \theta_{M3}$, whereupon indeed $\varphi_{Bk} = \psi$ correctly. Furthermore,

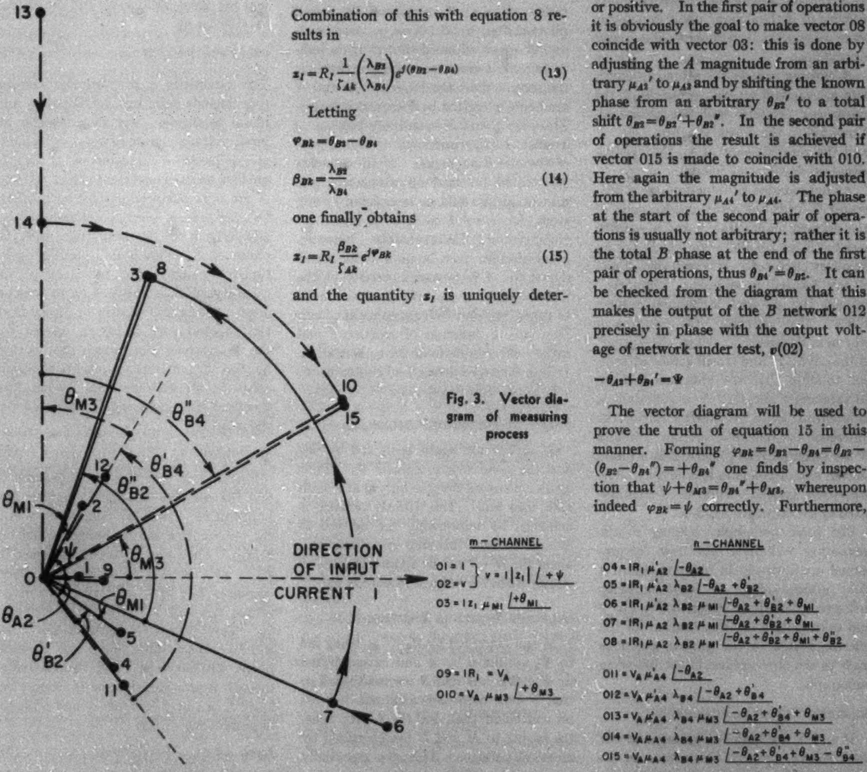

13
14
3
8
10
15
θ_{M3}
θ''_{B4}
θ'_{B4}
12
θ''_{B2}
θ_{M1}
2
ψ
θ_{M3}
1
9
0
θ_{M1}
DIRECTION OF INPUT CURRENT I
θ_{A2}
θ'_{B2}
5
4
11
7
6
m – CHANNEL
$01 = I$, $02 = v$ } $v = I|z_I| \angle +\psi$
$03 = Iz_I \mu_{M1} \angle +\theta_{M1}$
$09 = IR_I = V_A$
$010 = V_A \mu_{M3} \angle +\theta_{M3}$
n – CHANNEL
$04 = IR_I \mu'_{A2} \angle -\theta_{A2}$
$05 = IR_I \mu'_{A2} \lambda_{B2} \angle -\theta_{A2} + \theta'_{B2}$
$06 = IR_I \mu'_{A2} \lambda_{B2} \mu_{M1} \angle -\theta_{A2} + \theta'_{B2} + \theta_{M1}$
$07 = IR_I \mu_{A2} \lambda_{B2} \mu_{M1} \angle -\theta_{A2} + \theta'_{B2} + \theta_{M1}$
$08 = IR_I \mu_{A2} \lambda_{B2} \mu_{M1} \angle -\theta_{A2} + \theta'_{B2} + \theta_{M1} + \theta''_{B2}$
$011 = V_A \mu'_{A4} \angle -\theta_{A2}$
$012 = V_A \mu'_{A4} \lambda_{B4} \angle -\theta_{A2} + \theta'_{B4}$
$013 = V_A \mu'_{A4} \lambda_{B4} \mu_{M3} \angle -\theta_{A2} + \theta'_{B4} + \theta_{M3}$
$014 = V_A \mu_{A4} \lambda_{B4} \mu_{M3} \angle -\theta_{A2} + \theta'_{B4} + \theta_{M3}$
$015 = V_A \mu_{A4} \lambda_{B4} \mu_{M3} \angle -\theta_{A2} + \theta'_{B4} + \theta_{M3} - \theta''_{B4}$

Fig. 3. Vector diagram of measuring process

equation 12 demands that $(\theta_{A2}+\theta_{B4})=0$, i.e., in this case, since θ_{A2} is negative, $-\theta_{A2}+\theta_{B4}=0$. But $\theta_{B4}=\theta_{B4}'-\theta_{B4}''=\theta_{B2}'-\theta_{B2}''-\psi$, as was found in the foregoing; also, inspection of the diagram reveals that $\theta_{B2}'+\theta_{B2}''=\psi-(-\theta_{A2})$. Hence, indeed, $-\theta_{A2}+\psi+\theta_{A2}-\psi=0$. Equation 12 also requires that $\mu_{A4}\lambda_{B4}=1$. This relation is seen to be correct if one equates the magnitudes of vectors 010 and 015: $V_A\mu_{M3}=V_A\mu_{A4}\lambda_{B4}\mu_{M3}$. From equation 15 finally follows that

$$z_l=R_l\frac{1}{\zeta_{Ak}}\frac{\lambda_{B3}}{\lambda_{B4}}$$

Equating the magnitudes of vectors 03 and 08 gives $Iz_l\mu_{M1}=IR_l\mu_{A3}\lambda_{B2}\mu_{M1}$; Thus $z_l=R_l\mu_{A3}\lambda_{B3}$. From the relation $\mu_{A4}\ \lambda_{B4}=1$ it follows by definition that $\mu_{A3}\zeta_{Ak}=\mu_{A4}=1/\lambda_{B4}$, whereupon the required proposition is proved.

The measurement of the remaining two independent transmission quantities follows essentially the same pattern, and the pertinent relations will be merely stated without repetitious explanations. Referring to the measurement of the transfer admittance y_V and to Fig. 4, it is seen necessary to insert a known resistance r_t in series with z_t taking care that $r_t<<z_t$ so that the phase of the output current i is not altered appreciably: M is thus fed by ir_t. One has

Position 1 of S

$$ir_t\tau_{M1}+ir_t\underline{\tau_{N1}}=\delta_1\rightarrow 0 \qquad (16)$$

therefore

$$\tau_{M1}=-\tau_{N1} \qquad (17)$$

Position 2 of S

$$Vy_Vr_t\tau_{M1}+V\underline{\mu_{A2}}e^{j\theta_{A2}}\lambda_{B2}\underline{e^{j\theta_{B2}}}\tau_{N1}=\delta_2\rightarrow 0 \qquad (18)$$

therefore

$$y_V=-\frac{1}{r_t}\mu_{A2}\lambda_{B2}e^{j\theta_{A2}}e^{j\theta_{B2}} \qquad (19)$$

Position 3 of S

$$V\tau_{N3}+V\underline{\tau_{N3}}=\delta_3\rightarrow 0 \qquad (20)$$

therefore

$$\tau_{M3}=-\tau_{N3} \qquad (21)$$

Position 4 of S

$$V\tau_{M3}+V\underline{\mu_{A4}}e^{j\theta_{A4}}\lambda_{B4}\underline{e^{j\theta_{B4}}}\tau_{N3}=\delta_4\rightarrow 0 \qquad (22)$$

Since

$$\left.\begin{aligned}\mu_{A4}&=\mu_{A2}\zeta_{Ak}\\ \theta_{A4}&\equiv\theta_{A2}\end{aligned}\right\} \qquad (23)$$

$$1=\mu_{A2}\zeta_{Ak}\lambda_{B4}e^{j\theta_{A2}}e^{j\theta_{B4}} \qquad (24)$$

Therefore

$$y_V=-\frac{1}{r_t}\frac{\beta_{Bk}}{\zeta_{Ak}}e^{j\varphi_{Bk}} \qquad (25)$$

The final quantity, the transmission ratio ϱ_V, is measured exactly as y_V (see Fig. 4), except that M is fed by v instead of ir_t. It should suffice merely to state the result

$$\varrho_V=\frac{\beta_{Bk}}{\zeta_{Ak}}e^{j\varphi_{Bk}} \qquad (26)$$

As a matter of orientation it is well worth while to consider the orders of magnitude of the various quantities likely to occur. The phase aspect will be deleted for the moment and $\beta_{Bk}=1$ will be assumed. The voltage v_{PM} required at the input to the detector P will be assumed to be 0.1 volt which is supposed to guarantee measurement with prescribed accuracy. The gain of the N circuit will be assumed to equal that of the M circuit. Table I gives the assumed and the computed quantities for both active and passive X circuits and is believed to be self-explanatory since the identities are indicated for reference. It is seen that in these instances the gain of M would vary between 10 and 5,000 and that the A circuit would be required to vary between a gain of 500 and a loss of 0.2.

For the same three examples for X, and for certain assumed terminating impedances z_t, Table II gives the pertinent data for the determination of y_V. As is seen, in the passive case (the third) it is not

Fig. 4. Measurement of transfer admittance

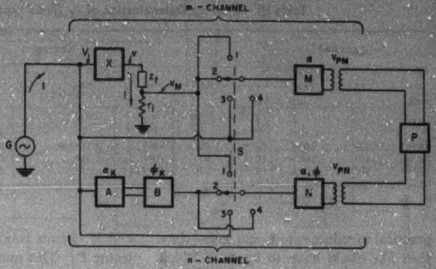


Table I. Data for Determination of z_l (Fig. 2)

Assumed					Computed								
v	μ_X	v_{PM}	z_l	R_l	V	μ_{M1}	I	V_A	μ_{A3}	μ_{M3}	μ_{A4}	ζ_{Ak}	z_l
volt		volt	ohm	ohm	volt		amperes	volt					ohm
					$\frac{v}{\mu_X}$	$\frac{v_{PM}}{v}$	$\frac{V}{z_l}$	$I\times R_l$	$\frac{v}{V_A}$	$\frac{v_{PM}}{V_A}$	1	$\frac{1}{\mu_{A3}}$	$\frac{R_l}{\zeta_{Ak}}$
10^{-2}	10^{0}	10^{-1}	5×10^{5}	10^{3}	10^{-2}	10	2×10^{-8}	2×10^{-5}	5×10^{2}	5×10^{3}	1	2×10^{-3}	0.5×10^{6}
10^{-2}	10^{2}	10^{-1}	10^{6}	10^{6}	10^{-4}	10	10^{-10}	10^{-4}	10^{2}	10^{3}	1	10^{-2}	10^{8}
10^{-4}	10^{-1}	10^{-1}	10^{4}	5×10^{2}	10^{-3}	10^{3}	10^{-7}	5×10^{-5}	2×10^{-1}	2×10^{3}	1	5	10^{2}

Table II. Data for Determination of y_V (Fig. 4)

Assumed					Computed								
v	μ_X	v_{PM}	z_t	r_t	V	i	v_M	μ_{M1}	μ_{A2}	μ_{M3}	μ_{A4}	ζ_{Ak}	y_V
volt		volt	ohm	ohm	volt	ampere	volt						ohm^{-1}
					$\frac{v}{\mu_X}$	$\frac{v}{z_t+r_t}$	$i\times r_t$	$\frac{v_{PM}}{v_M}$	$\frac{v_M}{V}$	$\frac{v_{PM}}{V}$	1	$\frac{1}{\mu_{A2}}$	$\frac{1}{r_t}\frac{1}{\zeta_{Ak}}$
10^{-2}	10^{3}	10^{-1}	10^{4}	10	10^{-5}	$\sim10^{-6}$	10^{-5}	10^{4}	1	10^{4}	1	1	10^{-1}
10^{-2}	10^{2}	10^{-1}	10^{4}	10^{2}	10^{-4}	$\sim10^{-6}$	10^{-4}	10^{3}	1	10^{3}	1	1	10^{-2}
10^{-4}	10^{-1}	10^{-1}	10^{2}	10	10^{-3}	$\sim10^{-6}$	10^{-5}	10^{4}	10^{-2}	10^{2}	1	10^{2}	10^{-3}

Table III. Data for Determination of ρ_V (Fig. 2 and 4)

Assumed			Computed						
v	μ_X	v_{PM}	V	μ_{M1}	μ_{A2}	μ_{M4}	μ_{A4}	ζ_{Ak}	ρ_V
volt		volt	volt						
			$\frac{v}{\mu_X}$	$\frac{v_{PM}}{v}$	$\frac{v}{V}$	$\frac{v_{PM}}{V}$	1	$\frac{1}{\mu_{A2}}$	$\frac{1}{\zeta_{Ak}}$
10^{-2}	10^{6}	10^{-1}	10^{-8}	10	10^{6}	10^{7}	1	10^{-6}	10^{6}
10^{-2}	10^{2}	10^{-1}	10^{-4}	10	10^{2}	10^{3}	1	10^{-2}	10^{2}
10^{-6}	10^{-1}	10^{-1}	10^{-5}	10^{5}	10^{-1}	10^{4}	1	10	10^{-1}

practical to reduce r_t to 1 ohm or less since then μ_{M1} would have to be $>10^4$ which might be hard to realize for a broad frequency band. It is better to be satisfied with a fair approximation in this borderline case.

For the sake of completeness, Table III is prepared as an aid to the determination of ρ_V, which in this case is identical to μ_X. As a cross check one observes that in every instance equations 1 and 2 are satisfied, giving identities for z_d and z_t respectively.

As was pointed out in connection with Figs. 2 and 4 it will be noted that each measurement involves adjustment of phase and amplitude in the n channel. This implies a significant property of the ideal detector, namely, to be sensitive to amplitude, phase, and direction of unbalance,[7] so that it may be termed a polarized directional detector. Also, it must be of the differential type in order to perform the operation of indicating a null when two equal signals are fed into it.

The "approach to zero," as it occurs in ordinary rectifiers, will be ruled out when balancing to within $\pm 10^{-6}$-volt input is the goal. Rather the "transit through zero," as it occurs in d-c Wheatstone bridges, must take place.

The necessity to demand all the sharpness of null detection possible to achieve will not seem unwarranted if detection is to be maintained for unfavorable signal-to-noise ratios. For that reason it will be imperative to incorporate in the response a large amount of "affinity" to the measuring frequency and, preferably, to no other, i.e., to demand a high degree of selectivity for any single frequency. The performance of the polarized rectifier is fully described elsewhere[23] and will be assumed known for the purposes of this discussion.

The Polarized Rectifier and Associated Apparatus as a Null Detector

Figs. 2 and 4 have already revealed the differential character of the measurement, for they showed outputs from the N and M circuits being combined in the detector P. This matter will now be pursued more carefully by studying Fig. 5. Consider a vector **A** at a certain time instant and let it rotate with the angular frequency ω such that the end point remains on the circle about O and of radius A. Assume another vector **B** which may be varied in amplitude and phase until $\mathbf{B}+\mathbf{A}=O$. Let x be an axis through the origin, the direction being arbitrary, and let **A** and x enclose an angle φ. Suppose that it is possible to observe the quantity $A \cos \varphi = A_x$. (This, but for a factor of proportionality, is precisely the quantity which the polarized rectifier[23] does indicate.) The locus of the end point of **B** must be on the circle just described; another locus is a line orthogonal to the x axis at a distance $-A_x$ from O. One way to observe **B** would be to observe B and B_x, but it is not a unique way since **B** might be at I or, equally well, at II. The orthogonal locus is not sufficient to define the vector **B**, since there exist infinitely many vectors satisfying the specified requirement, for example, vector **B′**. Another condition must be imposed.

Suppose another arbitrary direction is possible simultaneously. The second preferred direction will be assumed orthogonal to the x direction, although no requirement of orthogonality is stipulated. Thus the y axis is established, and it will be assumed that the component $A_y = A \cos(90-\varphi) = A \sin \varphi$ is also observable. It is now clear that the end point of **B** must also lie on a line orthogonal to the y axis at a distance $-A_y$ from O and **B″** represents one of the infinitely many vectors satisfying this condition. When both observations are used simultaneously, the adjustment of **B** is accomplished uniquely, rapidly, and unambiguously with the final criterion $\mathbf{B}=-\mathbf{A}$.

It is obvious that if **A** approaches orthogonality to one of the axes, the component along that axis approaches zero. Thus the cancelling effect of the vector **A** upon **B** (in the output) is lost and the galvanometer deflection in the direction for which, say, the **A**-component vanishes may become very large. To prevent this, provision should be made to rotate the coordinate system x,y by 45 degrees into another system x',y'. Since, of course, the process of balancing one vector against another is a process invariant with the system of co-ordinates, the final results are identical in either case.

The question may be raised why this differential type of detection is being advocated. The answer is that a comparison type of method, in which **A** is first ob-

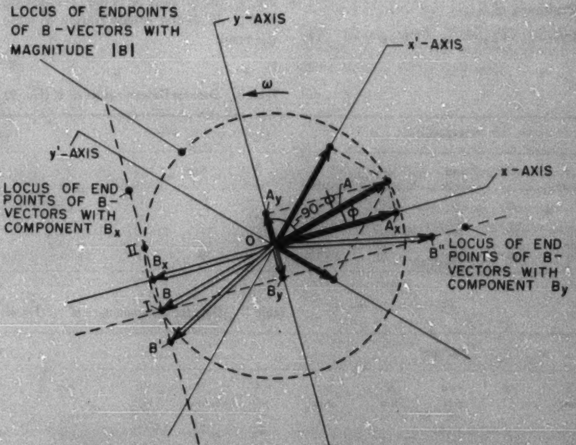

LOCUS OF ENDPOINTS OF B-VECTORS WITH MAGNITUDE |B|
y-AXIS
x'-AXIS
ω
y'-AXIS
x-AXIS
LOCUS OF END POINTS OF B-VECTORS WITH COMPONENT Bx
A
Ay
90-φ
φ
Ax
O
B''
LOCUS OF END POINTS OF B-VECTORS WITH COMPONENT By
II
Bx
B
I
By
B'

Fig. 5. Vector diagram for polarized rectifier as a null detector

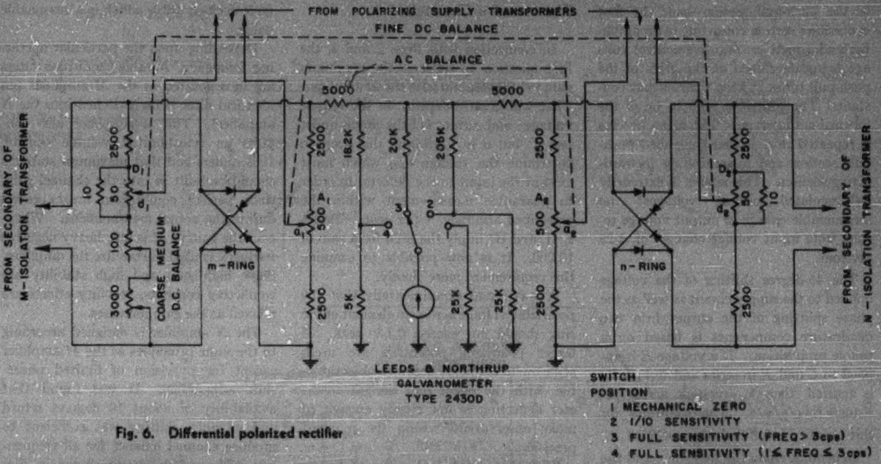

FROM POLARIZING SUPPLY TRANSFORMERS
FINE DC BALANCE
AC BALANCE
5000
5000
FROM SECONDARY OF M-ISOLATION TRANSFORMER
2500
10
50
100
3000
D_1
d_1
COARSE MEDIUM D.C. BALANCE
m-RING
2500
500
2500
A_1
a_1
18.2K
5K
20K
205K
25K
25K
1
2
3
4
2500
500
2500
A_2
a_2
n-RING
D_2
d_2
2500
50
10
2500
FROM SECONDARY OF N-ISOLATION TRANSFORMER
LEEDS & NORTHRUP GALVANOMETER TYPE 2430D
SWITCH POSITION
1 MECHANICAL ZERO
2 1/10 SENSITIVITY
3 FULL SENSITIVITY (FREQ > 3cps)
4 FULL SENSITIVITY (1 ≤ FREQ ≤ 3cps)

Fig. 6. Differential polarized rectifier

served and then **B** is established such that **B**=**A**, is considered to represent an inherent source of error, since memory must be depended upon to keep four data: direction and magnitude of A_x and A_y. Besides, no significant simplification in design, would result.

If in the general situation illustrated in Fig. 5 the vectors **A** and **B** are thought of as transfer factors τ_{M1} and τ_{N1} (see, for example, equation 5) and the co-ordinates x and y are visualized as the directions of the polarizing voltage vectors, if further it is understood that the observation of B_x+A_x is made by a galvanometer deflection δ_x and that of B_y+A_y is made by a deflection δ_y of a second galvanometer, then the specific situation of the present polarized rectifier detector system is at hand and the condition of vector identity manifests itself by a double null $\delta_x=\delta_y=0$.

Since the null indicator (a d-c galvanometer of the d'Arsonval type) forms an integral part of the polarized rectifier, representing the device capable of integration, over a number of cycles, and since this is the device which must suppress all a-c components, its a-c behavior must be considered. The galvanometer used was the Leeds and Northrup type *2430D* with the constants: undamped natural period=3.6 seconds; critical damping resistance=25,000 ohms; d-c sensitivity 4×10^{-10} amperes (amp) per millimeter (mm); d-c coil resistance=485 ohms.

Experimental work results in the recommendation to damp the galvanometer critically for all frequencies >3 cycles, and to use aperiodic damping (5,000-ohm) for frequencies between 1 and 3 cycles.

Attention is called to the order of magnitude of null detection. If the voltage phasor passing through the m channel (see Fig. 2) causes the galvanometer to deflect, say, 50 mm, then the adjustment of the phasor through the n channel should be within ±1/2 mm, if a precision of 1 per cent is to be maintained. For such precisions and using the varistor element mentioned in reference 23, the semiconductor rings have to be well-thermostated (at 50 degrees centigrade) to avoid slow galvanometer drifts caused by temperature variations. If selenium or germanium is used, thermostating may not be required.

The schematic diagram of one of the differential polarized rectifiers, say P_x, is shown in Fig. 6. A second polarized rectifier, P_y, being fed by the y polarization and being otherwise a duplicate of the one shown in Fig. 6, should be imagined to co-exist.

For a differential type of rectifier four independent adjustments would be required, comprising the "polarization null". (This is an outcome of the polarized rectifier study.[24]) Fortunately it was found experimentally that both the d-c and the a-c balance may be carried out simultaneously for the two sections, if only the control potentiometers are operated properly. Referring to Fig. 6 and realizing that the a-c signals from the n and m channels are fed in-phase to the rectifier, the required polarity shift is brought about by the manner in which the rings are interconnected. Thus assuming d_1 to be plus with respect to a_1 (because of polarizing voltage) and d_2 plus with respect to a_2, alternate pairs of elements in the m and n rings are being made conductive simultaneously with the effect that the d-c currents in the galvanometer branch oppose each other (this is the reason why vectors **A** and **B** in Fig. 5 have been drawn in opposite quadrants). Now, the d-c "fine" balance is so arranged that when resistance d_1D_1 decreases, resistance d_2D_2 increases. A similar statement applies to the a-c balance.

A 4-position switch for the galvanometer is used: In position 1, no energy is fed to the galvanometer; it is merely shunted by its critical damping resistance. In this position the mechanical zero is adjusted or checked. In position 2, the sensitivity is approximately one-tenth full sensitivity. Position 3 is used for frequencies $f\geq3$ cycles and position 4 for $1\leq f<3$ cycles (with aperiodic damping and approximately the same d-c sensitivity).

Another complication results from the observed phenomenon that the d-c galvanometer deflection is as sensitive to changes of polarizing voltage as it is to those of signal voltage: after a polarizing balance, bothersome galvanometer drifts occur, caused by the changes just mentioned. The situation demands automatic regulation of polarizing supply: it must be of high quality but does not need to cover a large range nor does its speed of response need to be fast. It must, however, extend down to 1 cycle. The polarizing supply is shown in block form in Fig. 7.

The push-pull stages employ *6BA6* remote cut-off pentodes. The input signal to the push-pull section feeds the first controlling section consisting of an amplifier and a rectifier, negative control voltage being developed at the grids of the push-pull tubes to keep their output constant. The control is seen to be of the forward-acting type. The same process is repeated in a second controlled push-pull section and a controlling forward-acting section. This results in extremely well-regulated output voltages: no measureable change in output voltage results from input voltage changes of ±25 per cent.

The 45-degree shifting of the voltage applied to the supply circuit as well as the phase splitting of the output into two quadrature components is based on a simple proposition. If a voltage E is applied to C and R in series and if $V_C = V_R$ is assured, then $\boldsymbol{V}_R$ = voltage across resistor $= E/\sqrt{2}\underline{/\pi/4}$. The condition for this equality is fulfilled when $\omega CR = 1$.

Depending upon the frequency and the choice of C, the correct R is adjusted in terms of calibrated dial readings of an appropriate helipot. In order to make the amplitude at 45 degrees the same as that at 0 degrees it is necessary to take care that the potential dividing circuit is such that $R_2/(R_1+R_2) = 1/\sqrt{2}$. For the phase splitter, the same components of C and R as for the phase shifter are used, the relation $\boldsymbol{V}_R = \boldsymbol{V}_C\underline{/+\pi/2}$ is here also guaranteed for $\omega CR = 1$ and adjustment is accomplished simultaneously by ganged helipots and selector switches.

The *m* and *n* Channels

In connection with Figs. 2 and 4, the first question to settle is the maximum gain to be incorporated in the M amplifier. That will greatly depend on the output voltage v and current i of the circuit under test X, but it is possible at this stage to determine the voltage v_{PM} which must exist at the input to the detector in order to guarantee measurement within the prescribed accuracy, 1 per cent. Tables I, II, and III imply the need for a gain of 10,000. It is now possible to examine this requirement more closely.

It was found experimentally that the polarizing voltage across an element of the ring should not exceed 0.1-v peak. If larger polarizing potentials are used, semiconductor noise becomes objectionable, with the result that the galvanometer deflection is not steady enough to make observations within the required precision.

With this in mind, it will next be assumed that detection of $\pm 1/2$ mm is possible (a conservative assumption). Then, to detect variations with an error of 1 per cent, the minimum deflection should be 50 mm. With the help of the analysis available for the polarized rectifier,[23] it is then possible to deduce the requirement for v_{PM}; namely, for a galvanometer sensitivity of 4×10^{-10} ampere/mm. v_{PM} need not be larger than 0.0364-v peak. Since it may be assumed that v in Figs. 2 and 4 is $\geq 10 \times 10^{-6}$-volt, the gain of the M amplifier need not exceed about 3,500, or 68.5 decibels (db), which is a reasonable value.

Depending upon the particular operating frequency, suitable band-pass filters are incorporated in the M amplifier (an identical filter being switched into the N amplifier). The M amplifier also contains an uncalibrated volume control. It is understood that maximum stability should be built in into the channel gain, since perfect equilibrium must prevail during the actual measurement. Whenever the gain of M is low, heavy negative feedback is used to stabilize the amplifier. Since high gain and high stability are conflicting concepts, stability efforts are relaxed as the gain increases.

The N amplifier is designed according to the same principles as the M amplifier except for provision of limited phase-shift adjustment. It was found that availability of about 10 degrees retard and advance shifting was sufficient to produce channel balance for all frequencies. The simple circuit shown in Fig. 8(A) is used. C_2 is a 10-turn variable air capacitor, while R_1, C_1, and R_2 are selected by means of switches. Because of the required shift at frequencies as low as 1 cycle, R_1 and R_2 are high-megohm resistors, the maximum for R_1 being 2×10^9 ohms. For that reason the shifter must be followed by an electrometer tube stage. The transmission ratio for this shifter is

$$\varrho = \frac{V}{E} = \frac{1}{1 + \dfrac{R_1}{R_2}\dfrac{(1 + j\omega C_2 R_2)}{(1 + j\omega C_1 R_1)}} \qquad (27)$$

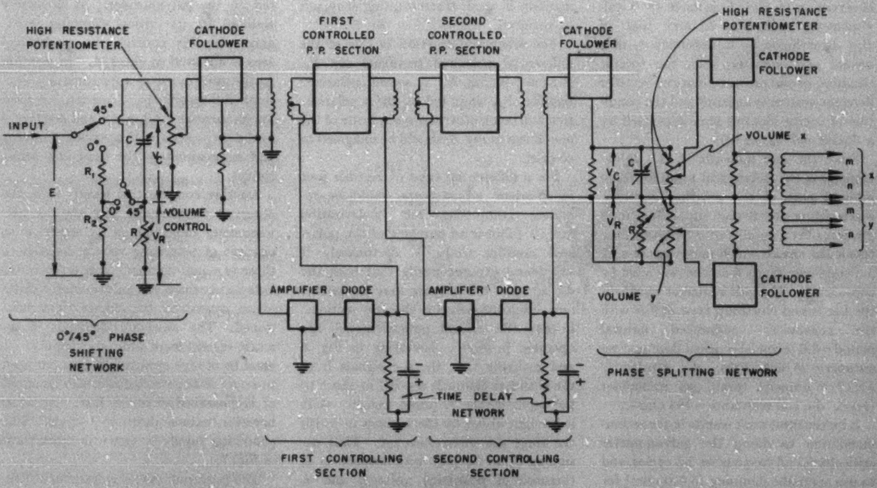


Fig. 7. Polarizing supply circuit

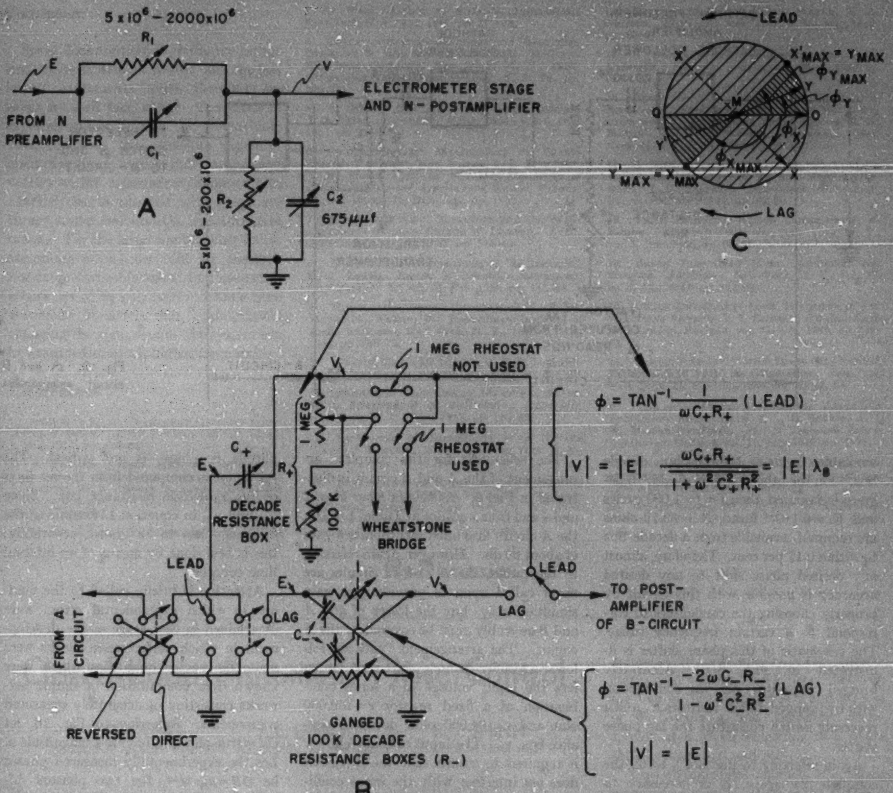

5 x 10^6 − 2000x10^6
R1
E
V
FROM N PREAMPLIFIER
C1
ELECTROMETER STAGE AND N − POSTAMPLIFIER
.5x10^6 − 200x10^6
R2
C2 675μμf
A
LEAD
X'
X'MAX = YMAX
ϕYMAX
Y
ϕY
Q
M
O
ϕX
Y'
ϕXMAX
Y'MAX = XMAX
X
LAG
C
1 MEG RHEOSTAT NOT USED
V1
1 MEG
E1
C+
R+
DECADE RESISTANCE BOX
100 K
1 MEG RHEOSTAT USED
WHEATSTONE BRIDGE
ϕ = TAN⁻¹ 1/(ωC+R+) (LEAD)
|V| = |E| ωC+R+/(1 + ω²C+²R+²) = |E| λB
LEAD
FROM A CIRCUIT
LAG
E1
V1
C−
LAG
LEAD
TO POST-AMPLIFIER OF B-CIRCUIT
REVERSED
DIRECT
GANGED 100K DECADE RESISTANCE BOXES (R−)
ϕ = TAN⁻¹ −2ωC−R−/(1 − ω²C−²R−²) (LAG)
|V| = |E|
B

Fig. 8. Phase shifters

A—Limited phase shifter for N circuit B—Full phase shifter, B network C—Vector diagram for B network

and the phase shift is

$$\varphi = \tan^{-1} \frac{(R_1/R_2)(\omega)(C_1R_1 - C_2R_2)}{(1+\omega^2C_1^2R_1^2)+\frac{R_1}{R_2}(1+\omega^2C_1C_2R_1R_2)} \tag{28}$$

It is seen from equation 28 that both retarding and advancing shifts are possible, depending on the relative magnitudes of the time constants C_1R_1 and C_2R_2. The circuit is so arranged that $(R_1/R_2) = 10 =$ constant, and it is therefore obvious that lagging phase shift ($C_1 = 0$) is always accompanied by a loss of more than 20 db. For this reason a postamplifier has to be incorporated in the N amplifier to boost the gain to the maximum decided for the M amplifier, 68.5 db. Amplitude of N is adjusted by means of a 15-turn 200,000-ohm helipot.

The full phase shifter (i.e., of the 360-degree type) for the B-circuit should ideally produce phase displacement independent of frequency. Although such a shifter has been deviced[24] and is also effective in the more elaborate phase-measuring systems,[2,8] it was deemed advisable to forego this advantage and to compute the phase displacement for greater accuracy. The present design centers about the well-known capacitor-resistor phase-shift lattice shown in Fig. 8(B): the pertinent equations are derived without difficulty and are indicated in the figure.

The magnitude of the output voltage is completely unattenuated, and the phase shift is always lagging. While $\varphi = 0$ and $\varphi = -180$ degrees are always possible by making $R_- = 0$ and $R_- = \infty$, respectively, a continuous approach $\varphi \rightarrow 0$ and $\varphi \rightarrow -180$ degrees is not possible for the latter case and is possible only with difficulties for the former one. Therefore, one generally has to be satisfied with attaining $\varphi_{\min} > 0$ and $\varphi_{\max} < 180$ degrees. Referring to Fig. 8(C), let $MO = E$ and let the phase shifter introduce a lagging angle φ_x bringing the output V in position MX. By reversing the polarity of E, MO goes over into MQ and MX into MX' so that the phase becomes $-(\varphi_x + \pi)$. Assuming that $\varphi_{\max}$ brings the vector into position $X_{\max}$, all points within sectors $OMX_{\max}$ and $QMX_{\max}'$ are made available by the lattice. Assuming $\varphi_{\max} = -160$ degrees, $f = 1$ cycle, and $C_- = 10\times^{-6}$ farad, there sults $R_- = 90{,}400$ ohms which is still

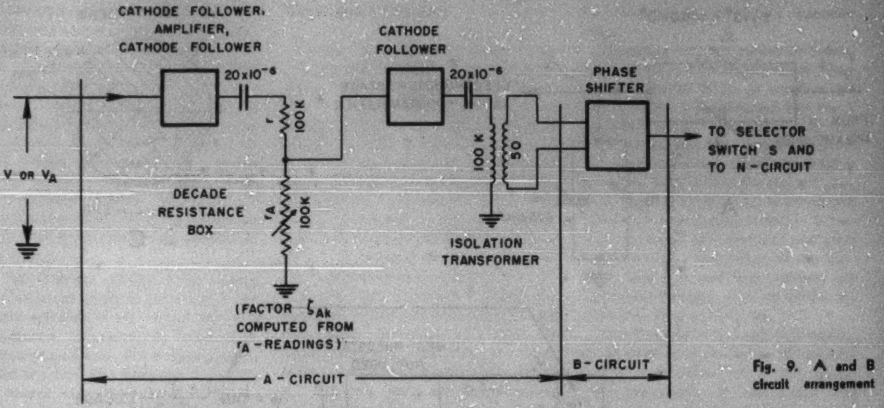


Fig. 9. A and B circuit arrangement

available from a 100,000-ohm decade box. On the other hand, if $\varphi_{\min}=-5$ degrees is assumed, then for $f=1{,}000$ cycles and $C_-=0.110^{-6}$ farad, $R_-=69.12$ ohms are required, available from a decade box to within 0.17 per cent. Therefore, almost any desired phase shift to any desired accuracy is possible with this circuit by properly choosing the condensers to correspond to a certain frequency range. The schematic of this phase shifter is illustrated in Fig. 8(B). Accurately bridged capacitor C_- in conjunction with two ganged decade resistors R_- and a reversing switch constitute the lag phase shifter.

Again referring to Fig. 8(C), to fill the complete 2π space, it is necessary to change networks and to provide for region $OMY_{\max}$. If V within that region has a position MY so that the network produces a leading phase shift φ_y, then, by polarity reversal, MY' is also available. Thus regions $OMY_{\max}=OMX_{\max}'$ and $QMY_{\max}'=QMX_{\max}$ are added and the full 2π phase shifter is realized. The network used for leading phase is shown in Fig. 8(B). Unfortunately, phase displacement is accompanied by magnitude change. For $\varphi_{\max}=20$ degrees and for $f=1$ cycle and $C_+=10\times10^{-6}$ farad, $R_+=43{,}800$ ohms—easily available from a decade resistor. For $f=1{,}000$ cycles and $C_+=0.1\times10^{-6}$ farad, $R_+=4{,}380$ ohms—equally convenient. For low leading phase shift the lowest frequency $f=1$ cycle is the governing factor. At that frequency, the 1-megohm rheostat has to be inserted, making possible the establishment of a phase shift of ~0.92-degree lead. (The value of the resistance is bridged.)

The schematic for the complete arrangement of the A and B circuit is illustrated in Fig. 9. As is seen from Figs. 2 and 4 and from a glance at Tables I to III, the A circuit also must incorporate a gain of about 60 db. However, again referring to the tables, the A and N circuits are never called upon to provide high gain simultaneously, but the losses of the A and B assembly must be covered by the A circuit. The arrangement "cathode follower-amplifier-cathode follower" transfers the input voltage to a series combination of a fixed resistor $r=100{,}000$ ohm and a 100,000 ohm decade resistance box, r_A. The input cathode follower is required to assure that the A circuit does not interfere with the input conditions (for very high resistors R_f, Fig. 2, it becomes necessary to go over to an electrometer input stage for A). The output cathode follower presents a low source impedance to the voltage divider: if this precaution is omitted, stray capacitances begin to introduce errors at relatively low frequencies. The divided voltage is fed to another cathode follower which is needed as an impedance-matching device to transfer energy to the isolation transformer with the low secondary impedance (50-ohm): the latter is important to guarantee accurately computable phase shifts from the B settings. Notice the large (20×10^{-6} farad) blocking condensers for transmission down to 1 cycle. To compensate for all voltage losses, a stabilized additional gain is built into the A amplifier.

Nonambiguous balancing of the M and N circuits is usually possible in two pairs of adjustment: the first choice of the component to be adjusted, i.e., whether amplitude or phase, is not critical. The ζ_{Ak} factor is computed from the readings of the r_A-decade resistance box. Since β_{Bk} and φ_{Bk} in equation 15 contain ω, the frequency has to be known accurately: this is best done by means of an interval time counter.

Attention is briefly called to the manner in which experimental errors were determined in that stage of the development in which performance checks were made for vector quantities which were known from computation of simple networks consisting of accurately measured components. Referring to Fig. 10, let OA = true phasor u having a magnitude u. Let the experimentally measured phasor be $OB=u_m e^{-j\Delta\varphi}$, the two phasors differing by a phasor whose magnitude is Δu. Then, considering the differential triangle ABC, one has

$$\Delta u^2=(u_m-u)^2+u^2(\Delta\varphi)^2$$

Thus the error $\epsilon=\Delta u/u$ becomes

$$\epsilon=\left[\left(\frac{u_m}{u}-1\right)^2+(\Delta\varphi)^2\right]^{1/2} \tag{29}$$

an expression which contains both the error in magnitude and that in phase.

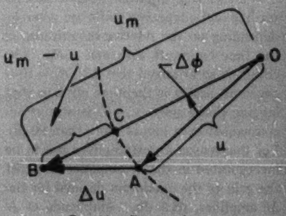


Fig. 10. Error evaluation

Conclusion

Space does not permit giving the rather lengthy measurement results, and the following statements must therefore be taken at their face value. Considerable experimental work with the system outlined indicates that absolute measurement of network transfer and transmission vectors within a frequency range of 1 to 1,000 cycles is possible with great accuracy under unfavorable signal-to-noise ratios. For the latter approaching unity, accuracies of ±1 per cent are possible. For more favorable signal-to-noise ratios accuracies of ±1/10 per cent have been obtained. It is felt that, appropriately changing the components, the system can be extended to much higher frequencies.

References

1. Accurate Radio-Frequency Microvoltages, **Myron C. Selby.** *AIEE Transactions*, vol. 72, pt. I, May 1953, pp. 158–64.

2. A Precise Direct Reading Phase and Transmission Measuring System for Video Frequency, **D. A. Alsberg. D. Leed.** *Bell System Technical Journal*, New York, N. Y., vol. 28, Apr. 1949, pp. 221–38.

3. Direct Reading Phasemeter, **L. H. O'Neill, J. L. West.** *Review of Scientific Instruments*, New York, N. Y., Apr. 1950, pp. 271–73.

4. Electronics for Communication Engineers (book), ed. by **John Markus, Vin Zeluff.** McGraw-Hill Book Company, New York, N. Y., 1952, "Measuring Vector Relationships," **Y. P. Yu**, pp. 269–72.

5. A Direct Reading Phase Shift Meter, **J. D. Eisler.** *Geophysics*, Tulsa, Okla., vol. 6, Oct. 1941, pp. 311–17.

6. Properties and Applications of Phase-Shifted Rectified Sine Waves, **J. T. Tykociner, L. R. Bloom.** *Bulletin No. 339, University of Illinois Engineering Experimental Station*, Urbana, Ill., vol. 40, no. 15, Dec. 1942, pp. 1–52.

7. Electronic Null Detectors for Use With Impedance Bridges, **Horatio W. Lamson.** *AIEE Transactions*, vol. 66, 1947, pp. 535–40.

8. A Wide Band Phasemeter, **J. R. Ragazzini, L. A. Zadeh.** *Review of Scientific Instruments*, New York, N. Y., vol. 21, Feb. 1950, pp. 145–48.

9. Modulated Beam Cathode Ray Phase Meter, **A. Watton, Jr.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 32, May 1944, pp. 268–72.

10. Vector-Response Indicator, **B. D. Loughlin.** *AIEE Transactions*, vol. 59, June 1940, pp. 355–57.

11. Simplified Vectorscope Measures Phase, **W. L. Firestone, R. A. Richardson.** *Electronics*, New York, N. Y., Sept. 1953, pp. 180–83.

12. Impedance Measurements with Square Waves, **F. Rockett.** *Electronics*, New York, N. Y., Sept. 1944, pp. 138–40, 336, 338.

13. Electronics for Communication Engineers (book), ed. by **John Markus, Vin Zeluff.** McGraw-Hill Book Company, New York, N. Y., 1952, "Measuring Phase at Audio and Ultrasonic Frequencies," **E. R. Kretzmer**, pp. 264–69.

14. Electronic Phase Angle Meter, **E. L. Ginzton.** *Electronics*, New York, N. Y., May 1942, pp. 60–61.

15. A Gain and Phase Measuring Set, **F. B. Anderson.** *Electrical Engineering*, Mar. 1953, p. 245.

16. Sensitive Measurement of Impedance, Reactance, Frequency and Phase Relationships with an Electronic Phase Bridge, **R. H. Brown.** *Review of Scientific Instruments*, New York, N. Y., vol. 13, 1942, pp. 277–81.

17. A New Method of Determining Voltage and Phase Relations in an A.C. Bridge Network, **J. R. Barnhart.** *Instruments*, Pittsburgh, Pa., vol. 14, Apr. 1941, pp. 89–90.

18. Phase Sensitive Bridge Detector, **P. H. Hunter.** *Electronic Industries*, New York, N. Y., vol. 5, June 1946, pp. 60–61.

19. Phase Indicating Null Indicator for Bridges. *Electronics*, New York, N. Y., vol. 17, no. 8, Aug. 1944, pp. 242–60.

20. Phase Detectors—Some Theoretical and Practical Aspects, **L. I. Farren.** *Wireless Engineer*, London, England, vol. 23, Dec. 1946, pp. 330–40.

21. Direct Reading Vector Impedance Bridge, **Technical Staff VFB Line Corporation.** *Tele-Tech*, New York, N. Y., June 1949, pp. 40–42, 64.

22. The ϕ-Detector, A Detector Valve for Frequency Modulation, **J. L. H. Jonker, A. J. W. M. van Overbeek.** *Philips Technical Review*, Eindhoven, Holland, vol. 11, July 1949, pp. 1–11.

23. The Ring Modulator as a Polarized Rectifier, **A. J. Hermont.** To be presented at 1955 AIEE Winter General Meeting.

24. Phase Shifting Up to 360°, **F. A. Everest.** *Electronics*, New York, N. Y., Nov. 1941, pp. 46–49, 122.

No Discussion

Closed-Loop Automatic Phase Control

P. F. ORDUNG
ASSOCIATE MEMBER AIEE

J. E. GIBSON
ASSOCIATE MEMBER AIEE

B. J. SHINN
STUDENT MEMBER AIEE

Synopsis: The dynamics of a closed-loop method of phase control of a local oscillator are discussed in this paper. An equivalent circuit is derived, on which further analysis of the system is based. While a linearized system and a system which uses time domain filtering have appeared in the literature, this paper presents a new, simple, and direct approach to the optimization of the system. A more general analysis is given and the general requirements for the nonlinear element and the subtraction device are pointed out. A method of design for the equalization filter is given. The theory is verified with experimental results.

Closed-Loop Control of the Phase of an Oscillator

IN DEVICES such as the sweep circuits of a television receiver and in pulse code modulation receivers, it is necessary in order to have successful operation, that not only the frequency but also the phase of a local oscillator be controlled by the incoming signal. Of several possible methods, perhaps the most flexible and the one with the best performance is that which uses a closed-loop system of control of the oscillator. In this method, the phase of the incoming signal is compared to the phase of the local oscillator by means of a phase detector. The output from the phase detector is used to control the frequency of the local oscillator in such a manner that the difference between the phase of the incoming signal and the phase of the locally generated wave is maintained constant in time; i.e., the frequency of the local oscillator is synchronized with the frequency of the incoming signal. The dynamics of the closed-loop system of phase control of the local oscillator are discussed in this paper.

Fundamental Theory

A block diagram of a typical closed-loop system for phase control of an oscillator is shown in Fig. 1. The incoming signal is $f(t)$. The local oscillator generates the voltages $g(t)$ and $g(t-\tau)$. The parameter τ is a constant time delay between the two output voltages $g(t)$ and $g(t-\tau)$ of the oscillator and therefore corresponds at a given frequency to a phase shift between $g(t)$ and $g(t-\tau)$.

The two multiplier channels, when their respective outputs $f(t)g(t)$ and $f(t)g(t-\tau)$ are subtracted, provide the instrumentation of the phase detector. This may be seen as follows: A multiplier may be regarded as a balanced modulator. If two sinusoidal voltages of different frequencies are fed into a balanced modulator, the output consists of the sum of two sinusoidal voltages, the frequency of one being the sum and the other being the difference of the two input frequencies. If ω_1 and ω_2 are the two input angular fre-

---

Paper 54-268, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–24, 1954. Manuscript submitted March 24, 1954; made available for printing May 4, 1954.

P. F. Ordung, J. E. Gibson, and B. J. Shinn are all with the Dunham Laboratory of Electrical Engineering, Yale University, New Haven, Conn.

This paper is the result of research on the synchronizing of pulse code modulation receivers, carried on under the support of the United States Army Signal Corps contract no. *DA-36-039-SC-5416*, and is a digest of Technical Report no. 9, July 1953, of that contract.

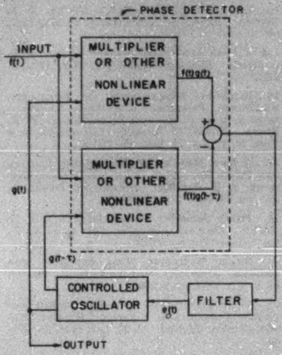


Fig. 1. Block diagram of a closed-loop system for the control of the phase of an oscillator

quencies, the instantaneous value of the voltage of difference frequency in the output is proportional to sin $[(\omega_1-\omega_2)t+\theta]$ where θ is a constant. Since $[(\omega_1-\omega_2)t+\theta]$ is the instantaneous phase difference between the two input voltages, the instantaneous amplitude of the voltage of difference frequency at the output of the modulator is seen to be a sinusoidal function of the instantaneous difference in the phase between the two input voltages. A single multiplier is thus a phase detector. (The foregoing argument suggests that any device into which two frequencies are fed and which is capable of producing a voltage of difference frequency in its output could be used to replace the multiplier. Thus the multiplier might be replaced by a peak detector, a square-law detector, an average detector, etc.)

When the outputs of the two multiplier channels shown in Fig. 1 are subtracted, the algebraic sign of the amplitude of the low-frequency component of voltage obtained at the subtraction point depends upon whether the phase of $f(t)$ is advancing or retarding relative to that of $g(t)$. This is a fundamental requirement for the comparison device. This property may be seen most easily in Fig. 2, in which the input wave $f(t)$ is saw-toothed and in which the two locally generated waves $g(t)$ and $g(t-\tau)$ are square pulse trains. For the phase relations shown in Fig. 2, the average of the output $e_1(t)$ of the first multiplier exceeds the average of the output $e_2(t)$ of the other multiplier. Thus the average of the voltage output at the subtraction point (see Fig. 1) is positive. Had $g(t-\tau)$ rather than $g(t)$ been more nearly in phase with $f(t)$, the average of the output at the subtraction point would have been negative. This dependence of sign of the average of the output at the subtraction point upon the direction of shift of phase of $g(t)$ relative to $f(t)$ provides the necessary discriminator action needed for control of the phase of the local oscillator so that the oscillator may be made to change in such a way that the average of the output at the subtraction point is reduced to zero. (The foregoing example illustrates one property of the types of waves that may be employed in this device. If the local oscillator voltage is shifted in time by an amount t' from the phasing that would yield null output at the subtraction point, the waves used must be such that the voltage $f(t)g(t-t')-f(t)g(t-t'-\tau)$ is an odd function of t'. Thus if t' is positive the average output is positive, and if t' is negative the average output is negative.)

The voltage at the subtraction point in Fig. 1 consists of the superposition of two components. One is a function of the instantaneous difference of phase between $f(t)$ and $g(t)$ and is therefore the component desired for control of the frequency of the oscillator. The other is at a high frequency since it is a function of the instantaneous sum of the phases of $f(t)$ and $g(t)$. One function of the low-pass filter provided is that only the desired, or low-frequency, component of voltage at the subtraction point may influence the tuning of the oscillator.

The filter has the added function of controlling the relative immunity of the system to disturbance from random noise that may be superimposed upon $f(t)$. This may be seen as follows: Suppose the filter is a simple resistance-capacitance (R-C) integrating circuit with a relatively long time constant. The output at any instant of time t of an R-C integrating circuit consists of a weighted average of the input from time t to approximately one time constant earlier. If the time constant of the filter is sufficiently long, voltage at the subtraction point arising from random noise superimposed upon

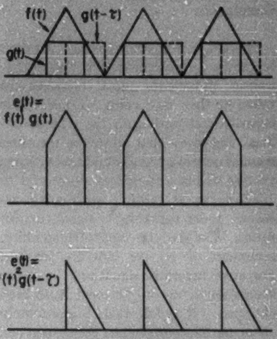


Fig. 2. Outputs of the multiplier channels with nonsinusoidal inputs

$f(t)$ will largely average to zero at the output of the filter; whereas the voltage component that is a function of the difference in the phase of $f(t)$ and $g(t)$ will have a finite value. Hence the longer the time constant of the filter, the greater will be the immunity of the system to disturbance from any random noise that may be superimposed upon $f(t)$.[1]

If $W(t)$ is the response of the filter to a unit impulse input, the output voltage $e_0(t)$ of the filter may be expressed in terms of its input $f(t)g(t)-f(t)g(t-\tau)$ as follows

$$e_0(t)=\int_{-\infty}^{t}[f(x)g(x)-f(x)g(x-\tau)]W(x)dx$$

$$e_0(t)=\int_{-\infty}^{t}f(x)g(x)W(x)dx-\int_{-\infty}^{t}f(x)g(x-\tau)W(x)dx$$

Each of the integrals in the second expression of this equation has the character of a short-time cross correlation. One property of cross correlation is that as the correlation time, (the time constant of the R-C filter), is increased, the contribution in the output of the filter resulting from random noise superimposed upon $f(t)$

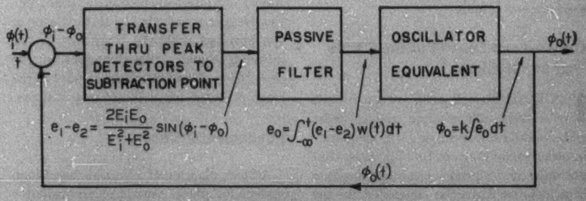

$\phi_i(t)$
t
$\phi_i-\phi_0$
TRANSFER THRU PEAK DETECTORS TO SUBTRACTION POINT
PASSIVE FILTER
OSCILLATOR EQUIVALENT
$\phi_0(t)$
$e_1-e_2=\frac{2E_iE_0}{E_i^2+E_0^2}\sin(\phi_i-\phi_0)$
$e_0=\int_{-\infty}^{t}(e_1-e_2)w(t)dt$
$\phi_0=k\int e_0dt$
$\phi_0(t)$

Fig. 3. Phase equivalent circuit of a closed-loop system for the control of the phase of an oscillator

tends to vanish, as compared to the contribution resulting from $f(t)g(t)$ and $f(t)g(t-\tau)$.[1]

The output of the filter is used to control the tuning of the local oscillator. Perhaps one of the easiest ways to achieve control of the tuning is to shunt the tank circuit of the oscillator with a reactance tube. Upon the grid that controls the amount of reactance offered by the reactance tube is impressed the voltage output from the filter. If the phasing of the system is correct, any tendency of the oscillator to deviate from the frequency of the incoming signal will result in the appearance at the output of the filter of a voltage of the appropriate sign so as to cause the reactance tube to retune the oscillator and correct the deviation.

Equivalent Circuit

The equivalent circuit, subsequently to be obtained, although derived for the case in which peak detectors are used (as in Fig. 1) nevertheless applies for all instrumentations. In the equivalent circuit shown in Fig. 3, the only difference encountered when different instrumentations of Fig. 1 are used is in the analytical expression of the transfer through the phase detector portion.

For the purpose of derivation of the equivalent circuit, the input voltage $f(t)$ and the oscillator voltages $g(t)$ and $g(t-\tau)$ will be expressed by the following equations.

$$f(t)=E_i \sin\left[\omega_0 t+\phi_i(t)\right] \tag{1}$$

$$g(t)=E_0 \sin\left[\omega_0 t+\phi_0(t)\right] \tag{2}$$

$$g(t-\tau)=E_0 \sin\left(\omega_0 t+\phi_0(t)-\omega_0\tau\right) \tag{3}$$

In the foregoing equations, $\phi_i(t)$ and $\phi_0(t)$ are phase angles that are functions of time which vary slowly relative to $\omega_0 t$. The time delay τ in the left side of equation 3 is replaced in the right side by the equivalent shift in phase of $\omega_0\tau$ radians at a frequency ω_0. The angular frequency ω_0 is the frequency of the voltage generated by the oscillator in the absence of any tuning effect exerted by the loop.

The equation for the transfer through a pair of ideal peak detectors to the subtraction point in Fig. 1 will now be obtained. The input to one of the peak detectors is $f(t)+g(t)$ and to the other is $f(t)+g(t-\tau)$. The output of a peak detector is a slowly varying voltage wave that passes through each of the successive peaks of the input wave. The output obtained at the subtraction point in Fig. 1 is the difference of the outputs of the two detectors. The output obtained at the subtraction point may therefore be found by taking the difference of the equations for the envelopes of the waves input to the two detectors. When the foregoing analysis is performed, the following equation for the output at the subtraction point is obtained.

$$e_1-e_2=\sqrt{E_i^2+E_0^2+2E_iE_0\cos(\phi_i-\phi_0)}-\sqrt{E_i^2+E_0^2+2E_iE_0\cos(\phi_i-\phi_0+\omega_0\tau)} \tag{4}$$

If each of the radicals in equation 4 is expanded into a series by means of a binomial expansion, and the first two terms of each of the series is retained, and if in the result a trigonometric substitution is made, the following approximate expression for the output at the subtraction point is obtained.

$$e_1-e_2=\frac{2E_iE_0}{\sqrt{E_i^2+E_0^2}}\sin\frac{\omega_0\tau}{2}\sin\left[\phi_i-\phi_0\right] \tag{5}$$

The greatest error resulting from the use of the binomial approximation occurs when $E_i=E_0$. The omission of the higher ordered terms in the expansion consists primarily of omission of the harmonic content in the output of the peak detector. This omission is relatively unimportant because of the nature of the closed loop in the system. In the case usually encountered in practice, one of the voltages is much larger than the other, and the error resulting from the use of the binomial approximation is negligible.

Equation 5 shows that the output obtained at the subtraction point is proportional to the sine of the instantaneous phase difference between $f(t)$ and $g(t)$. Thus the portion of Fig. 1 containing the peak detectors and the subtraction point may be represented as in Fig. 3 as an equivalent phase discriminator circuit,

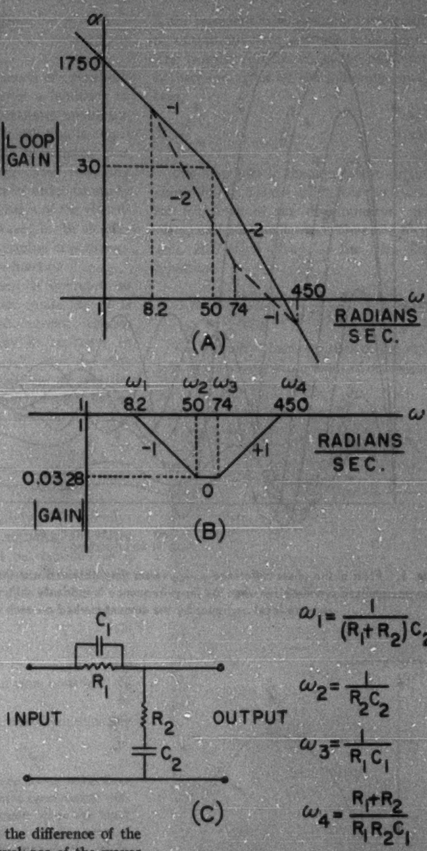


Fig. 4. Attenuation versus frequency diagram and schematic diagram of equalization filter required in Fig. 4(A)

A—Bode diagram of unequalized and equalized system
B—Required equalizer for Bode diagram above
C—Schematic for equalization filter

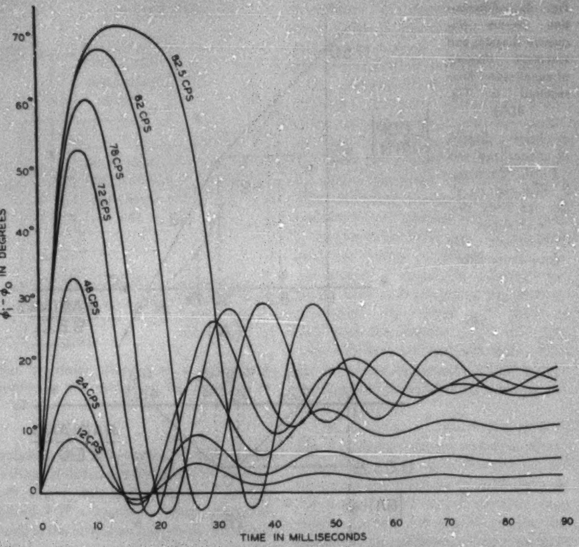


Fig. 5. Plots of the phase difference $\phi_i-\phi_0$ versus time obtained with the computer analogue of the unequalized synchronizer when the input frequency is suddenly shifted from that of the free-running local oscillator by the amount marked on each curve

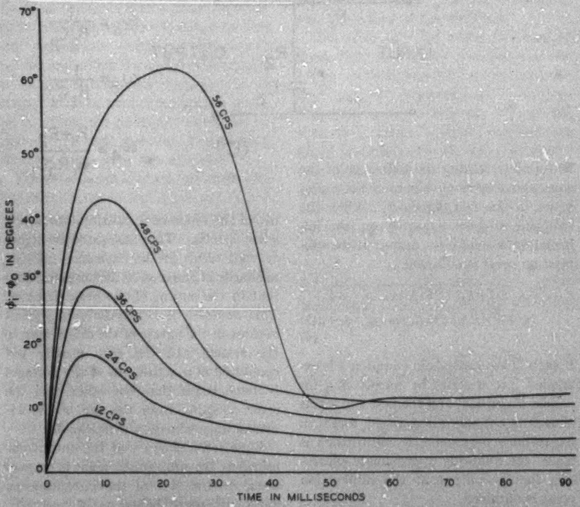


Fig. 6. Plots of the phase difference $\phi_i-\phi_0$ versus time obtained with the computer analogue of the equalized synchronizer when the input frequency is suddenly shifted from that of the free-running local oscillator by the amount marked on each curve. Note the rounding of the largest curve caused by nonlinearity of the system

which has as its input ϕ_i and ϕ_0 and as its output e_1-e_2. In order that as much output as possible be obtained for a given phase difference $\phi_i-\phi_0$, the time delay τ should be chosen so that $\sin \omega_0\tau/2$ is unity. Throughout the remainder of the paper τ will be regarded as so chosen.

The output obtained at the subtraction point in Fig. 1 is filtered by means of a passive filter and then is used to control the frequency of the local oscillator. In the equivalent circuit shown in Fig. 3, the filter is represented by its weighting function $W(t)$.

The oscillator shown in Fig. 1 will be assumed to be tuned by a reactance tube that is shunted across the oscillator tank circuit. The deviation $\delta\omega$ in angular frequency of the oscillator from its free-running angular frequency ω_0 will be assumed to be a linear function of the voltage impressed on the control grid of the reactance tube. Thus $\delta\omega$ is related to the output voltage e_0 of the filter by the following equation.

$$\delta\omega = ke_0 \tag{6}$$

In this equation, k is a constant of proportionality. The amount of the deviation of the phase of the oscillator when its frequency is deviated is denoted in equations 2 and 3 by ϕ_0. Since phase is the time integral of frequency, ϕ_0 is related to $\delta\omega$ by the following equation.

$$\phi_0 = \int \delta\omega dt \tag{7}$$

The substitution of equation 6 into equation 7 yields the following expression for the relationship of ϕ_0 to e_0.

$$\phi_0 = k\int e_0 dt \tag{8}$$

The frequency-controlled oscillator may therefore be represented in an equivalent circuit as in Fig. 3 by a box, the transfer through it given by equation 8.

The equivalent circuit for the phase control of an oscillator by means of a closed loop may be seen from Fig. 3 to be a single loop feedback structure that contains one nonlinear element and two linear elements. The input to the equivalent circuit is ϕ_i and the output is ϕ_0.

Steady State

In the steady state, the oscillator is synchronized to the frequency of the incoming signal, and $\phi_i-\phi_0$ is a constant. To find the steady-state value of $\phi_i-\phi_0$, let the angular frequency of the incoming signal be ω_i and that of the free-running oscillator to be ω_0. So that the oscillator may be synchronized to the frequency of the incoming signal, the voltage on the grid of the reactance tube that tunes the oscillator must have a value such that the frequency of the oscillator is deviated from ω_0 by an amount $\omega_i-\omega_0$. The substitution of $\omega_i-\omega_0$ for $\delta\omega$ in equation 6 yields the following expression from which the voltage on the grid of the reactance tube may be obtained.

$$\omega_i-\omega_0=ke_0 \tag{9}$$

The filter in the system is of the low-pass type; hence its d-c transfer may be taken as unity. Consequently e_0 in equation 9 may be replaced with the right side of equation 5. When the resulting expression is rearranged, the following equation is obtained. (As pointed out earlier, $\sin \omega_0\tau/2$ is taken as unity throughout this paper.)

$$\sin(\phi_i-\phi_0)=\frac{\omega_i-\omega_0}{\dfrac{2kE_0E_i}{\sqrt{E_i^2+E_0^2}}} \tag{10}$$

The value of $\phi_i-\phi_0$ when the oscillator is synchronized to the frequency of the incoming signal is obtained by solution of equation 10. It will be noted that no solution is possible and the oscillator cannot be synchronized when

$$\frac{\omega_i-\omega_0}{\dfrac{2kE_0E_i}{\sqrt{E_i^2+E_0^2}}}>1.$$

Furthermore for a given value of $\omega_i-\omega_0$, the phase difference $\phi_i-\phi_0$ tends to zero as the loop gain $2kE_iE_0/\sqrt{E_i^2+E_0^2}$ of the equivalent circuit (shown in Fig. 3) is increased.

Steady-State Stability Characteristics

In the steady-state mode of operation, the oscillator may exhibit a tendency to hunt so that $\phi_i-\phi_0$ fluctuates about the value obtained from the solution of equation 10. It is desirable to suppress this tendency for hunting. The following analysis is undertaken in order to study the stability characteristics of the closed-loop system when operated in the steady-state mode and to determine how to suppress any tendency for hunting.

Normally the system is operated so that in the steady-state mode, $\phi_i-\phi_0$ is much smaller than 90 degrees. Consequently $\sin(\phi_i-\phi_0)$ may be replaced in equation 5 with $(\phi_i-\phi_0)$. For the purposes of this analysis, the phase discriminator may therefore be regarded as a linear device, which has the following transfer function H_1.[2]

$$H_1=\frac{2E_iE_0}{\sqrt{E_i^2+E_0^2}} \tag{11}$$

For a preliminary study of the steady-state stability characteristics, the filter shown in Fig. 3 will be regarded as a simple R-C low pass type, which has the transfer function H_2.

$$H_2=\frac{1}{Ts+1} \tag{12}$$

H_2 is expressed in the LaPlace transform notation, and T is the time constant RC of the filter.

As seen from Fig. 3, and from equation 8, the transfer of the oscillator from input to output involves a single integration. The transfer functon H_3 of the oscillator is therefore given by the following equation.

$$H_3=\frac{k}{s} \tag{13}$$

The loop gain of the equivalent circuit shown in Fig. 3 is the product of the transfer functions of the discriminator, the filter, and the oscillator. Thus the loop gain $H_1H_2H_3$ is given by the following equation.

$$H_1H_2H_3=\frac{2kE_0E_i}{\sqrt{E_i^2+E_0^2}}\,\frac{1}{s(Ts+1)} \tag{14}$$

The stability characteristics of the system in the steady-state mode of operation

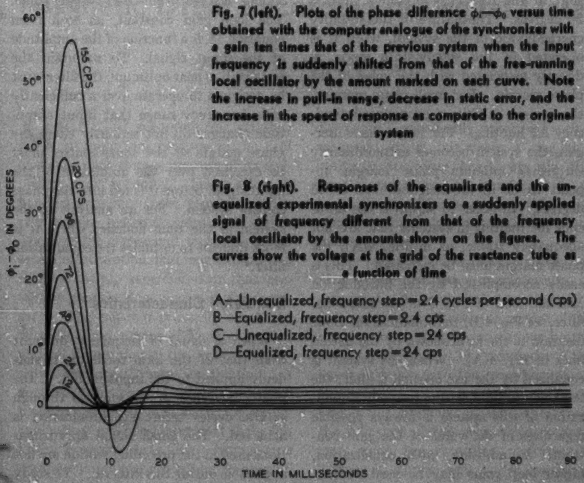


Fig. 7 (left). Plots of the phase difference $\phi_i-\phi_0$ versus time obtained with the computer analogue of the synchronizer with a gain ten times that of the previous system when the input frequency is suddenly shifted from that of the free-running local oscillator by the amount marked on each curve. Note the increase in pull-in range, decrease in static error, and the increase in the speed of response as compared to the original system

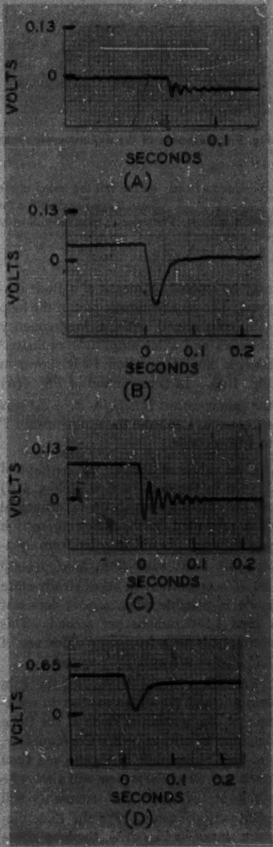


Fig. 8 (right). Responses of the equalized and the unequalized experimental synchronizers to a suddenly applied signal of frequency different from that of the frequency local oscillator by the amounts shown on the figures. The curves show the voltage at the grid of the reactance tube as a function of time

A—Unequalized, frequency step = 2.4 cycles per second (cps)
B—Equalized, frequency step = 2.4 cps
C—Unequalized, frequency step = 24 cps
D—Equalized, frequency step = 24 cps

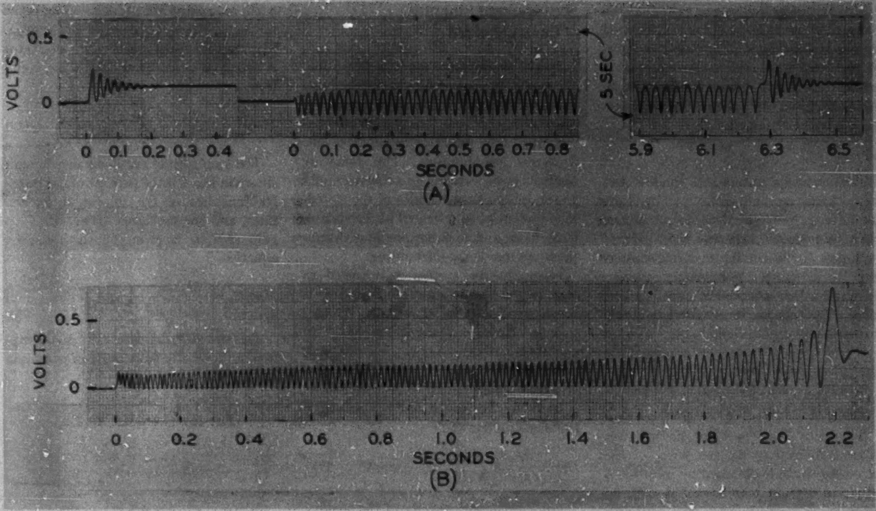


Fig. 9. **Response of the experimental synchronizer (both equalized and unequalized) for the largest values of $\omega_i-\omega_0$ at which lock-in may occur. $2kE_iE_0/\sqrt{E_i{}^2+E_0{}^2}=1,750$**

A—Unequalized; on the left the initial phase was such that synchronization occurred immediately. On the right the system went into a false mode where it remained until a random perturbation caused lock-in. $\omega_i-\omega_0=38$ cps$\times 2\pi=239$ radians per second

B—Equalized. Plate shows gradual pull-in of the system. $\omega_i-\omega_0=93$ cps$\times 2\pi=586$ radians per second

may be studied by means of a Bode diagram; i.e., an asymptotic plot of the loop gain constructed with a log-frequency abscisses and a log-amplitude ordinate.[3] Such a plot of equation 14 is shown in Fig. 4(A). In construction of Fig. 4(A) the gain constant $2kE_0E_i/\sqrt{E_i{}^2+E_0{}^2}$ was set equal to 1,750 and the time constant T to 0.02 seconds.

(The values of $2kE_0E_i/\sqrt{E_i{}^2+E_0{}^2}$ and T used in the construction of Fig. 4(A) are the values for which the experimental data presented in the following section was obtained. It will be noted from equation 10 that with $2kE_0E_i/\sqrt{{}_i{}^2E+E_0{}^2}=1{,}750$ operation in the steady-state mode is possible when $\omega_i-\omega_0$ does not exceed 1,750 radians per second. This corresponds to a frequency difference of $1{,}750/2\pi$ or 278 cycles per second.)

Regardless of the choices of the gain constant and the time constant in equation 14, the maximum slope of the Bode diagram is always -2. Consequently, the loop phase shift is always less than 180 degrees and the system will always be stable in the sense that transients will eventually die away. For the Bode diagram shown in Fig. 4(A), the loop phase shift at the frequency of unity gain may be shown to be -154 degrees. The phase margin, which is the difference between 180 degrees and the loop phase shift at the frequency of unity gain, is 26 degrees. The nature of the stability characteristic of the system, when it is operated in the steady-state mode, is directly associated with the value of the phase margin. For small phase margins, the system exhibits a pronounced tendency for hunting. For large phase margins, the system becomes extraordinarily sluggish. Optimum phase margin appears to be in the neighborhood of 45 degrees. In order to suppress the tendency for hunting in systems that are not adequately stable, as in Fig. 4(A), the phase margin must be increased. This is easily accomplished by the introduction into the loop of Fig. 3 of an additional filter, see Fig. 4(B) that modifies the Bode diagram to the form shown by the dotted lines in Fig. 4(A). When the system is equalized by the use of such a filter, the tendency to hunt is suppressed. Equalization of this system is always possible regardless of the value of the gain constant. In addition, with equalization, higher loop gains may be used without tendency to hunt than would otherwise be possible.[4]

The equalization departs from optimum if the value of the loop gain constant should change from the design value. The loop gain constant, as seen from equation 14, is a function of the amplitude of the input signal. To maintain the equalization near optimum, the filter must be designed to operate over a sufficiently wide frequency range that input amplitude changes do not seriously effect the phase margin or else some control must be exercised over the amplitude of the input signal before it is fed into the phase control system. For an analysis of the system in the time domain, where it is more difficult to optimize the equalization filter.[4]

Simulated Characteristics

When a signal of frequency different from that of the local oscillator is suddenly applied to the input in Fig. 1, the phase difference $\phi_i-\phi_0$ may go through several cycles before synchronization is achieved. The small signal approximations used in the preceding section are not applicable during this interval. To study the behavior of the system in the initial synchronizing interval, the circuit shown in Fig. 3 was simulated on a REAC analog computer. The phase discriminator was realized in the computer by means of a resolver. Since the resolver had a rotation that was limited to ±180 degrees, the use of the computer was limited to the study of those initial synchronizing conditions for which the $\phi_i-\phi_0$ did not exceed 180 degrees. The study was performed on a system which did not use the equalization filter, on the same system when it was equalized, and on a system which in addition to being equalized had the loop gain increased by a factor of 10. The results obtained, which are shown in Figs. 5, 6, and 7, clearly demonstrate the substantial improvement obtained through the use of the equalization filter. The improvement is obtained not only in the suppression of the tendency to hunt in the steady-state mode of operation but in the duration of the initial synchronizing interval as well. An increase in the frequency range over which synchronization could be accomplished may also be noted.

Experimental Characteristics

The preceding theories were verified on an experimental model in which the phase discriminator was instrumented by means of peak detectors and in which the oscillator was controlled by means of a reactance tube. The loop gain constant $2kE_0E_i/\sqrt{E_0^2+E_i^2}$ was 1,750 and the time constant T of the filter was 0.02 seconds. The experimental model was tested with and without the use of the equalization filter, which is shown in Fig. 4. The experimental results obtained, which are shown in Figs. 8 and 9, clearly show the improvement obtained with the use of the equalization filter.

The use of the equalization filter has a further advantage. The frequency difference $\omega_i-\omega_0$ in which the unequalized system may be guaranteed to always synchronize is smaller than that which is predicted by means of equation 10. The reason is that near the limits of $\omega_i-\omega_0$ as predicted by means of equation 10, the unequalized synchronizer may, if the initial phase difference at the instant of application of the input signal is favorable, go into a stable oscillation. This is demonstrated in Fig. 9(A). In the left side of Fig. 9(A), the initial phase difference $\phi_i-\phi_0$ was unfavorable to the oscillation mode; whereas in the right side, the initial phase difference was favorable to the oscillation mode. The oscillation mode persisted until a random disturbance occurred that shocked the system into the steady-state mode. In the equalized system the frequency range $\omega_i-\omega_0$ in which synchronization always occurs is substantially extended relative to that of the unequalized system.

Conclusions

Proper equalization results in substantial improvement in the performance of the system. In the system which was tested, the synchronization time was reduced by a factor of 10 and the pull-in range was extended by a factor of 2. Furthermore, operation in the false mode was eliminated by the equalization. When properly equalized, synchronizers can be designed for higher loop gains than would otherwise be possible, and the benefits of still smaller phase errors and shorter pull-in terms realized.

References

1. Theory of Servomechanisms (book), **Hubert M. James, Nathaniel B. Nichols, Ralph S. Phillips**, McGraw-Hill Book Company, Inc., vol. 25, Radio Laboratories Series, chap. 6, 1947.

2. Analysis of Synchronizing Systems for Dot-Interlaced Color Television, **T. S. George**. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 39, Feb. 1951, pp. 124–31.

3. Servomechanisms and Regulating System Design (book), **Harold Chestnut, Robert W. Mayer**. John Wiley and Sons, Inc., New York, N. Y., 1951.

4. Color-Carrier Reference Phase Synchronization Accuracy in NTSC Color Television, **Donald Richman**. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 42, Jan. 1954, pp. 106–33.

No Discussion

Magnetic Drum Recording of Digital Data

ALBERT S. HOAGLAND
ASSOCIATE MEMBER AIEE

MAGNETIC recording has attained an important position in the field of digital data storage because of its simplicity and economy. Information is handled in binary form, requiring a discrimination between only two states. These states correspond to two opposite senses of saturation, and new information is then merely recorded over that previously contained, erasing being unnecessary. This paper is primarily concerned with the development of a useful mathematical description of the process of magnetic recording of binary information. A concept of this process is developed which furnishes an analytical basis for design evaluation. While a large part of the following is applicable to binary storage on a continuous magnetic medium in general, the paper is specifically concerned with a magnetic drum unit using a ring-type magnetic recording head, wherein the type of magnetization is referred to as longitudinal. This is at present the most common form for a magnetic drum memory system.

The paper will first present a brief review of the distinctive characteristics of a magnetic drum digital recording unit, followed by a qualitative discussion of the input-output relations for various binary coding techniques. The remainder of the paper is devoted to the recording geometry which is of paramount importance. The major emphasis of the paper is on the influence of the geometrical parameters upon the recording performance. The basic storage unit under consideration is a single magnetic head servicing its associated peripheral track.

Drum Storage Characteristics

A moving storage surface for binary information can be specified in terms of its capacity in binary digits (bits) and an inherent access time, which is taken here to be the recurrence interval for a given bit of information. For a rotating drum this interval is the time for one revolution. The merit of the storage surface may be measured by the ratio of capacity to

---

Paper 54-271, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 22, 1954; made available for printing May 4, 1954.

Albert S. Hoagland is with the University of California, Berkeley, Calif.

The author wishes to express his appreciation for the interest shown in this work by Prof. Paul L. Morton, University of California, Berkeley, Calif. The co-operation of the California Digital Computer Project, supported by the U.S. Office of Naval Research, particularly with regard to the use of its facilities, is gratefully acknowledged.

access time as it is advantageous to increase the capacity, decrease the access time, or both.

$$C/A=v/s=p$$

where

C = capacity of track, bits
A = inherent access time
v = surface velocity of magnetic storage layer
s = cell length or distance assigned to one bit on the magnetic surface in the direction of its motion
p = information transfer rate, bits per unit time

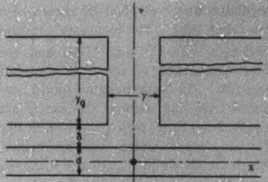
This ratio is then equal to the rate of information transfer which is an external measure of memory quality. A slight separation between the magnetic head and the magnetic surface allows an increase in orders of magnitude of the surface velocity over permissible contact values, more than overcoming the resulting loss in resolution (increase in the minimum value for s). Generally the time response will be more than adequate, particularly with the advent of ferrite heads. An increase in velocity involves mechanical considerations. A customary challenge then faced by the worker in the field of binary magnetic recording is that of effecting a reduction in the cell length.

The essential differences between this application and the magnetic recording of sound are the following:

1. Necessity for reliability as many hours of computation may be wasted by the misinterpretation of a single bit.
2. Two-level saturation recording on the magnetic surface.
3. Fundamental frequencies of operation, directly related to the bit rate, of the order of 100 kc per second.
4. Noncontact operation of the magnetic head which is synonomous here with high surface velocities (commonly of the order of 1,000 inches per second) and a narrower wave-length band. Wave length refers here to a sinusoidal variation of magnetization with distance along the storage surface.

Coding Techniques for Binary Storage

Considering the translation from a binary sequence to an appropriate recording current as a coding process, it is clear there is no unique choice. Binary storage coding techniques can be grouped broadly into two categories. One group comprises those methods which result in a useful output signal for every bit, while the second includes those codes which provide a useful output signal only upon a change in binary sequence or, equivalently, only for one of the two binary digits. The first group of codes is in a sense redundant in that the original binary sequence can be reconstructed, knowing only the first digit in addition to the changes in binary digit sequence. These divisions cannot be extended to uniquely characterize in every case the input current code. For any particular input current code the resulting output signal depends upon the recording system design and cannot be anticipated independently of this information. A brief description of the two most common codes follows:



d = THICKNESS OF MAGNETIC SURFACE LAYER
δ = SEPARATION DISTANCE BETWEEN THE MAGNETIC SURFACE AND THE MAGNETIC HEAD
γ = AIR GAP DISTANCE
y_g = AIR GAP DIMENSION NORMAL TO THE MAGNETIC LAYER
W (NOT SHOWN) = WIDTH OF THE MAGNETIC HEAD IN THE DIRECTION NORMAL TO THE x AND y COORDINATES

Fig. 1. Idealized magnetic head and surface geometry

1. Nonreturn-to-zero or NRZ method. In this method the input current when recording is continuous and a 1, e.g., is represented by a positive value of saturation current throughout the cell period and a 0 by a negative or reverse value of saturation current. Later the relation between input and output will be detailed but, for the present, as a crude first approximation, the output can be represented by the derivative of the associated input current. Then only changes in binary sequence are detected and this method falls uniquely into the second category described.
2. Return-to-zero or RZ method. Here a positive recording current pulse is applied for a 1 and a negative pulse for a 0 during the associated cell period. This method can therefore be employed to give an output signal for every bit. However, it also can be used when only changes in sequence are useful, the system being operated beyond the point at which the simple differentiating approximation for the output voltage is valid.

The choice of a particular code depends upon over-all reliability and digital processing requirements. In the remaining sections the word "writing" will be used to refer to the magnetization of the storage layer and the word "reading" will be used in the restricted sense of readback output voltage.

Influence of Geometry

This section considers the influence of the spatial parameters upon the recording behavior and their relation to the question of cell density. The following diagram indicates schematically the over-all storage transfer process involved in the magnetic recording of binary information. The peripheral co-ordinate is x

$$\underbrace{i(t)\rightarrow[M(x)-\phi_h(x)]}_{\text{writing}}\underbrace{\rightarrow\frac{vNd\phi_h}{dx}=e_0(vt)}_{\text{reading}}$$

where

$M(x)$ = distribution of magnetization on the surface
$\phi_h(x)$ = reading-coil flux as a function of the angular position of the drum
e_0 = open-circuit readback output voltage
N = number of reading turns

The input current is associated with the recorded pattern through the air-gap fringing field of the magnetic recording head. The output voltage can be directly related to a variation of core flux with drum angular position. To complete the transfer relation, the connection between the reading-coil flux and the surface magnetization $M(x)$ must be established.

The idealized recording system geometry used for the treatment of the spatial dependence is shown in Fig. 1. In the usual arrangement the following order of magnitudes holds: d, γ, and δ are of the order of several mils, while y_g and W are of the order of 50 times this. In view of the size of the area of interest, the air-gap region, relative to the surface curvature, it is reasonable to regard the storage layer as a sheet infinite in extent in the x direction. Further, the geometrical relations can be considered as if the problem were 2-dimensional. These idealizations are the basis of the analytical treatment of the recording geometry.

WRITING

The significance of the magnetizing field distribution will be discussed qualitatively. Suppose at some instant the air-gap field is at one of its writing saturation levels. Then for the symmetrical head structure assumed a maximum longitudinal magnetizing field exists throughout the surface layer at the plane $x=0$. This will effect a saturation condition in this layer cross section which is one of the two characteristic magnetic states that can be attributed to the surface medium with this mode of recording. By using saturation levels the prior magnetic history of this cross section is not important, and therefore the influence of the writing current at preceding instants can be neglected.

Due to the x extent of the fringing field, this cross section must undergo some finite displacement from the air-gap centerline before the magnetizing field is reversed, in order that this induced magnetic saturation state remain unchanged. This minimum displacement is essentially given by the effective writing field extent, as will be described.

Suppose the track is uniformly saturated in the negative x direction and the surface is stationary. The application of an oppositely directed saturation writing field will reorient the condition of magnetization in the immediate gap region and reverse the sense of saturation at the air-gap centerline. The effective writing field extent η is defined as the distance in the x direction from the air-gap centerline to a plane $x=\eta$, which is identified with one resulting boundary of the negatively saturated region.

The analysis is based on the following model. The applied magnetizing field is assumed to arise from a uniform magnetic surface charge density distribution on the pole faces of the magnetic head. This assumption corresponds to the uniform magnetization of the magnetic head core under an impressed input current. Further, as y_g is much greater than the other spatial parameters in this 2-dimensional structure, these pole faces are considered to extend to infinity in the positive y direction.

Using this model the writing field is identical to that arising from two semi-infinite charged planes. The horizontal writing field component H_x can easily be shown to be proportional to α; see Fig. 2. Based upon the hysteretic behavior of the surface material, η can be computed approximately as this distance may be related to a given fractional decrease in the horizontal field component from the reversing saturation level at $x=0$.

Experimentally this distance was measured in the following way. The track was initially saturated in the negative sense. Then with the drum stationary a positive saturation magnetomotive force was applied to the magnetic head and removed. The resulting surface magnetic state was observed optically with a microscope through the application of an extremely fine carbonyl iron powder in an oil suspension, the particles being drawn to the saturation transition boundaries. The correlation with the idealized model appears reasonably good in view of the nature of the problem and this simple model seems suitable for evaluating the relative influence of the system parameters in writing.

The writing definition η, representing the minimum distance between the two saturation states that can be achieved on the storage surface for a given design, establishes the minimum cell length limit due to writing. It should be noted that because of the nonlinear magnetic surface characteristic, necessary for storage, the writing definition is better than the sharpness of the magnetizing field considered alone warrants. This would not carry over to reading regarded from a principle of reciprocity.

When the input current is held constant at one of the saturation levels all surface cross sections traverse the same magnetostatic writing field, including a region sufficient to produce a common saturation condition. Thus a closely uniform magnetic state results, i.e., there is no variation with the peripheral co-ordinate. It will be assumed that $M=M_x$ and that the senses of saturation correspond to the positive and negative x directions. That the main orientation is longitudinal is in accord with experimental behavior and, indeed, would be expected to predominate where longitudinal recording is used. M_x then resembles in space the recording current in time, as long as a current step does not appreciably effect the previously recorded transition region as outlined earlier. Due to the insensitivity of H_x with y and for practical values for d, it is reasonable to neglect any variation with y.

READING

As indicated in the diagram of the recording process, to analytically formulate the transfer behavior a relation must be established between M_x and $\phi_h(x)$. With the low field levels on readback, the magnetic core acts closely as a linear element. Then the output for an arbitrary two-level input pattern can be synthesized by superposition, using the response to a step input. However, M_x for this step function is only qualitatively known, with no specific functional form. Thus the approach taken here is to consider the wave-length response in reading and use a filter concept, this filter operating upon the wavelength spectrum of M_x. Conventional filter concepts can then be employed in describing the transmission behavior of the system in reading, neglecting base-line effects. The wave-length bandwidth needed can be intimated from the writing definition. After this filter is postulated the over-all behavior to a step function input will be outlined.

For the idealized recording unit indicated, the derivation of the desired relation between the reading coil flux and M_x is straightforward. The various factors entering into this relation have been given and confirmed in the literature[1] and these results will be merely assembled here in the desired form; i.e.

$$\phi_h(\lambda)=KH(\lambda)M_x(\lambda)$$

where

K = a constant and includes the multiplying factor of head-width

$H(\lambda)$ = a function of the recorded wave length λ and the geometrical parameters

$$H(\lambda)=\left[\lambda\left(1-\epsilon^{-\frac{2\pi d}{\lambda}}\right)\right]\left(\epsilon^{-\frac{2\pi\delta}{\lambda}}\right)\left(\frac{\sin \pi\gamma/\lambda}{\pi\gamma/\lambda}\right)$$

These factors hold for the realizable

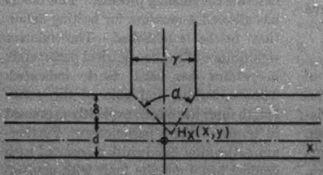


Fig. 2 (left). Geometrical definition of the angle α

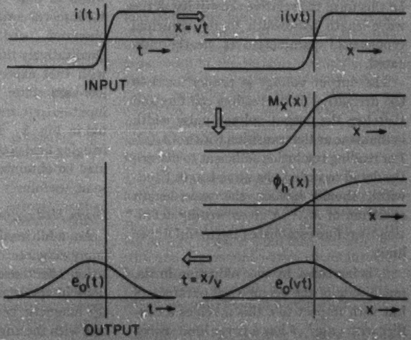


Fig. 3 (right). Input-output step function transfer process wave lengths encountered in this magnetic recording application. Because of the ideal differentiation operator in the overall transfer process there is no particular loss in generality in accepting the extension of $H(1/\lambda)=H(u)$ to the limit $u=0$. Then $H(u)$ can be regarded as the response of a low-pass wave-length filter with only attenuation. As u approaches zero, $H(u)$ approaches $2\pi d$. Then the gain of this filter is equal to $2\pi d$ or is directly proportional to the depth of the magnetic storage surface. Let the bandwidth of this filter be defined as that value of u at which the response is down 3 decibels, designating this point as $u=u_c$. An explicit expression for the bandwidth can be obtained by first replacing each factor of $H(u)$ by an approximation, satisfactorily representing the original term out to u_c; i.e., for $0\leq u\leq u_c$

$$H(u)\approx[2\pi d(1-\pi du)][1-2\pi\delta u][1-c(\pi\gamma u)]$$

where $0<c\leq 0.21$.

A linear approximation in u has been chosen for the air-gap factor. In this form c actually should be a function of the relative magnitude of γ compared to d and δ, $c=0.21$ corresponding to the entire attenuation arising from this term. In practice γ, d, and δ are usually of the same order of magnitude and the contribution of the air-gap factor represents only a fraction of the total attenuation at the 3-decibel point. For an improved approximation c should therefore be less than 0.21. It is apparent here that u_c will be fairly insensitive to the actual value assigned to c.

Taking the first-order terms in u

$$H(u)\approx H(0)[1-\pi du-2\pi\delta u-c(\pi\gamma u)]$$

where $H(0)=2\pi d$. Then

$$u_c\approx\frac{0.292}{\pi d+2\pi\delta+c\pi\gamma}$$

In this form the relative sensitivity of the wave-length bandwidth to the principal geometrical parameters is clearly indicated.

The output voltage is proportional to the derivative of the reading coil flux and therefore the output voltage pulse width is the same as the transition width of $\phi_h(x)$. For reading resolution sufficient to discern the detail recorded, the wave-length bandwidth should include the wave-length spectrum of M_x. A given writing definition therefore establishes the useful upper limit.

u_c is increased by any reduction in the magnitude of the geometrical parameters but with respect to d this involves a conflict with gain. δ has a lower limit owing to the necessity for noncontact operation. A restriction on a continued decrease of γ arises because of the effect of this upon writing, as eventually a point will be reached beyond which it will be impossible to maintain the saturation level of the air-gap fringing field at the storage medium. A lower limit for d appears in obtaining a satisfactory signal-to-noise ratio. This limit may arise in terms of a signal-to-surface noise ratio because of the difficulties of getting a uniform coating in a thin layer, particularly if a spray technique is used.

Where noncontact operation is essential it is seen that a definite upper limit exists for u_c. An optimum d may exist for a given separation distance, for as d is decreased eventually this will have little influence on increasing u_c while still resulting in a proportional gain reduction. The nonlinear characteristic of the storage surface favors writing.

Step Function Response

The correlation between the scales of distance along the surface and time is given by $x=vt$. However, even for instantaneous switching of the input current the change in magnetization M_x on the surface will have a finite gradient with a transition region equal to η, as indicated earlier. Thus a loss in short wave-length detail results when writing in terms of a 1-to-1 correspondence between the time and space domains. Reading of the recorded pattern has been shown equivalent to transmission through a low-pass wave-length filter which results in an attenuation of the shorter wave-length components of M_x. This low-pass filter action results in a loss of resolution in reading and the output voltage pulse, proportional to the derivative of the reading-coil flux, exhibits the customary broadening effect from such a low-pass transfer operator. Experimentally the output voltage pulse width in time is considerably more than the interval corresponding to the optically observed surface transition region. Qualitatively this pulse width is in accord with that expected on the basis of this low-pass filter concept. The over-all input-output transfer process is illustrated in Fig. 3. Due to the spread of the fringing field, a time advance phenomenon may be obtained. This effect may have some usefulness.[3]

Drum Modulation

An additional problem in the use of many magnetic drum memories is a modulation effect resulting from slight variations in the separation distance between the magnetic head and the magnetic surface with the angular position of the drum. This modulation effect can be formulated in the following manner in terms of the resulting maximum output voltage variation for a given wave length.

Let

δ_0 = mean value of the separation distance
ν = maximum deviation from the mean of the separation distance

then

$$\frac{|\Delta e_0|}{\bar{e}_0}=\frac{\epsilon^{-\frac{2\pi}{\lambda}(\delta_0-\nu)}-\epsilon^{-\frac{2\pi}{\lambda}(\delta_0)}}{\epsilon^{-\frac{2\pi}{\lambda}(\delta_0)}}$$

$$=\epsilon^{\frac{2\pi\nu}{\lambda}}-1$$

For $\nu<<\lambda$

$$\frac{|\Delta e_0|}{\bar{e}_0}=\frac{2\pi\nu}{\lambda}$$

It is seen that this modulation effect is independent of the mean separation distance and under the latter condition is directly proportional to the ratio of the maximum deviation to the recorded wave length. The recording field is not as sensitive to small changes in δ and hence "modulation" will not be relatively insignificant in writing. Modulation reduces the minimum-signal-to-maximum-surface noise ratio for the whole track circumference to less than that which holds for any given angular position. The foregoing result further yields a technique for the measurement and display of the eccentricity of rotating members which can be made extremely sensitive.

Conclusions

Most magnetic drum units operate at cell densities where considerable interference between the output response of adjacent input step functions is tolerated. The degree to which this is permissible depends somewhat upon the decoding scheme used, but the basic recording limitations arise from the transfer behavior outlined here. The importance of the geometry in the magnetic recording of binary information has been emphasized. A model allowing the analytical formulation of writing has been advanced and a wave-length filter concept developed to describe the reading process. This model has allowed a measure for writing definition to be established. The relative significance of the geometrical parameters in reading has been clearly indicated. The bandwidth and gain of the wave-length filter have been explicitly expressed in terms of these parameters. Furthermore, the filter bandwidth associated with reading has a definite limit with noncontact operation irrespective of any practical sacrifice in gain that may be possible. The nonlinear character of the storage surface is shown to result in an improvement of writing definition over the corresponding resolution in reading. Modulation effects have been shown to depend only upon the absolute variation in separation distance.

References

1. Reproduction of Magnetically Recorded Signals, R. L. Wallace. *Bell System Technical Journal*, New York, N. Y., vol. 30, Oct. 1951, pp. 1145–73.

2. Combined Reading and Writing on a Magnetic Drum, J. H. McGuigan. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 41, Oct. 1953, pp. 1438–43.

3. Magnetic Recording (book), S. J. Begun. Murray Hill Books, Inc., New York, N. Y., 1949.

◆

No Discussion

Electrical Protection of Telephone Systems

W. R. BULLARD
FELLOW AIEE

J. B. HAYS
NONMEMBER AIEE

H. O. SAUNDERS
ASSOCIATE MEMBER AIEE

OCCASIONALLY, telephone plant and equipment are subjected to voltages and currents from external sources which are very much larger than those used in normal operation. A reasonable reserve in dielectric strength and current capacity is provided, but to cope with the more severe extraneous voltages it is necessary to provide protective arrangements.

The two most common sources of these extraneous voltages are lightning and contacts with electrical power circuits. There may be other effects such as the static potentials left on open conductors by charged sand particles blowing through the air; or even the infrequent magnetic storms which have sizable earth currents associated with them and which produce potentials between widely separated points. However, experience over the years has shown that lightning and power are the major electrical disturbances for which protective arrangements must be designed.

Little can be done to control the frequency of lightning disturbances and the protective arrangements are designed solely to reduce the effects. On the other hand, the magnitude and duration of currents and voltages arising from contact between power and telephones systems, as well as their effects, may be controlled by protective arrangements in both systems. This latter problem has been the subject of joint study of engineers of the power and telephone industries over many years. Many joint reports have dealt with co-ordination of power and telephone construction or with protective arrangements applicable to joint or crossing poles. This paper outlines the principles of protection for subscriber stations and describes how these basic concepts fit into the plan of protection for an entire telephone system.

Purposes

Turning first to the purposes of protective arrangements, there are three primary purposes as follows:

1. To provide for the safety of persons, i.e., customers, workmen, and the public.
2. To minimize likelihood of damage to a customer's property.
3. To limit the damage to the telephone plant and minimize service interference.

Of these, the safety of people is of paramount concern.

Methods

The basic methods used to accomplish these purposes are as follows:

1. Provision of effective grounding of telephone lines at subscribers' premises when the lines are subjected to abnormal voltages.
2. Provision of suitable current paths between nominally grounded metallic structures and telephone wires, and limitation of voltages to a safe value.
3. Provision of co-ordinated electrical protection of power and telephone plant so that in the event of a contact the power circuit involved will be promptly and reliably de-energized.
4. Provision of paths to divert lightning currents into the ground.

It is very fortunate that, by the nature of things, these methods all work in the same direction to accomplish the primary purposes. In general, there is no conflict between them; hence one does not have to be compromised to aid the other.

General Principles of Protection

It is necessary in protection to remember that the significant voltage which must be minimized is the voltage between two conductive parts which a person could touch simultaneously or between parts which must be insulated from each other against dielectric failure. The voltage of a conductive part with respect to true ground is not necessarily significant unless one of the two parts mentioned remains essentially at true ground potential during the abnormal voltage conditions. The latter condition seldom applies in practice because a ground usually must be connected to earth at a remote point to represent true ground.

At the risk of appearing to explain the obvious, two examples of the difference from the safety standpoint between potential to ground and potential difference may be worth while.

1. A modern form of the "Faraday cage" is a steel-bodied automobile. If one should be so unfortunate as to have a power wire fall on his car, he will be perfectly safe as long as he stays inside (even though the car body may be energized at several thousand volts with respect to ground potential) because all of the surfaces are at the same voltage. But unless he is unusually agile he cannot get in or out of the car safely. To avoid the chance of being killed, he must jump out, or in, in such a manner that he never touches the car and the ground simultaneously.

2. Small birds are perfectly safe in flying around and perching on high voltage power wires because they can never touch more than one of the wires at a time but a bird big enough to touch more than one wire or a wire and a grounded object simultaneously is taking a terrific chance—as witnessed by the number of dead hawks, cranes, etc., which are pulled out of power wires.

While it is not possible to have the cus-

Paper 54-289, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 29, 1954; made available for printing May 4, 1954.

W. R. Bullard is with Ebasco Services, Inc., New York, N. Y., J. B. Hays is with the Bell Telephone Laboratories, Inc., Murray Hill, N. J., and H. O. Saunders is with the American Telephone and Telegraph Company, New York, N. Y.

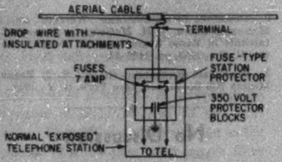

AERIAL CABLE
DROP WIRE WITH INSULATED ATTACHMENTS
TERMINAL
FUSES 7 AMP
FUSE-TYPE STATION PROTECTOR
350 VOLT PROTECTOR BLOCKS
NORMAL "EXPOSED" TELEPHONE STATION
TO TEL.

Fig. 1. Schematic of fuse-type protector formerly installed in cable areas

tomers or telephone workmen enclosed in a Faraday cage, a close and effective approach to it can be made. This approach can be summarized as follows:

1. To the extent that it is practicable, all wiring and metallic piping should be well grounded, preferably to the same grounding electrode.
2. On poles the telephone cable and plant should be bonded to effectively grounded power neutrals.

Since the telephone circuits cannot be connected directly to grounded metallic structures, it is necessary to make connections to ground through "protectors" to avoid disturbing normal telephone service. These protectors are simply devices which keep the circuit insulated in the absence of abnormal voltages and make direct connections in the presence of such voltages. These protectors, or lightning arrestors as they are sometimes called, usually take the form of low-voltage spark gaps, of which the most reliable and economical design for telephone purposes has been found to be two carbon electrodes with an air gap between them. For grounding noncurrent carrying parts, such as cable sheath, connection can usually be direct.

At the customer's premises the method consists of connecting the protective ground wires to the metallic piping systems or other suitable grounding electrode and preferably to the same electrode to which the neutral of the power system is connected. According to the National Electric Code, the customer's electric neutral and telephone protector ground must be connected to the metallic underground water system, public or private, when it is available. The preference for the water pipe as a telephone protector ground therefore comes as a result not only of its low impedance but also because of the common grounding it provides.

While the telephone facilities have been considered in the foregoing as the ones that are subject to energization, it must be recognized that it is also possible for other wire services, or even piping, to be energized by lightning or abnormal power conditions. If the latter condition exists, the telephone facilities provide a path away from the subscriber's premises rather than toward it. In either event, a common grounding electrode equalizes the potential differences within the premises.

On poles carrying both power wires and telephone cable, bonding can be accomplished by connecting at regular intervals the telephone cable sheath to the power system vertical ground wires associated with effectively grounded neutral wires. In addition to increased safety, bonds to neutrals (viz., vertical ground wires) will substantially limit the area over which cable troubles occur in the event of a contact or a lightning stroke.

Where telephone circuits are open wire they cannot be bonded directly to power neutrals or other nominally grounded objects because such direct connection would render them inoperative. However, they can be equipped with heavy-duty protectors (spark gaps with large electrodes) at intervals, and the grounding points of these protectors can be connected to power neutrals or other effective grounds.

To date such protectors have been used primarily where the power circuits with which contact may occur are in the 5-kv (or higher) voltage ranges. The protectors have breakdown voltages in the order of 3,000 volts. However, heavy-duty protectors having lower voltage breakdown (1,500–2,000 volts) are under development and it is expected that they will find increasing application.

Fuseless Station Protectors

One of the recent developments for exposed telephone stations in cable areas is a new heavier duty protector generally referred to as the "fuseless station protector."

Heretofore the type of telephone protection installed at a customer's premises was as shown in Fig. 1 and can be described as follows

1. A low-voltage spark gap operating at about 350-rms volts is provided from each wire to ground.
2. A fuse is provided in each line wire to interrupt the current if, and only if, it is large enough and persists long enough for the protector to overheat and start a fire.

The purpose of the fuse element in the Fig. 1 arrangement is to protect the protector and its associated wiring by limiting the currents that can flow through the protector under fault conditions. A blown fuse, however, does not remove the foreign potential from all parts of a telephone circuit, as the portion of the line side of the fuses may remain energized.

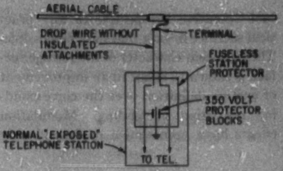

AERIAL CABLE
DROP WIRE WITHOUT INSULATED ATTACHMENTS
TERMINAL
FUSELESS STATION PROTECTOR
350 VOLT PROTECTOR BLOCKS
NORMAL "EXPOSED" TELEPHONE STATION
TO TEL.

Fig. 2. Schematic of fuseless protector presently installed in cable areas

This point is illustrated by the minimum and maximum current values which were used in the selection of the station fuse. A 7-ampere fuse was required so that it would not open under a 6.6-ampere fault current from the older constant current street-lighting circuits. A smaller fuse might operate and leave the high open-circuit voltage of the street-lighting circuit on the telephone conductors.

Further, each unopened ground connection provides a contribution to the current-carrying capacity of a grounding system or, conversely, the collective opening of many station fuses would reduce the ability of telephone plant to handle fault current. Thus, it can be reasoned that the fusible element should be as large as possible as long as it does not permit overheating of the station protector.

With these thoughts in mind, a heavier duty station protector has been developed and the new arrangement, which is now being used at telephones served by drop wire in telephone cable areas, is shown in Fig. 2.

Outwardly, the difference between Fig. 1 and Fig. 2 lies in the omission of the station fuse. There is, however, a fusible link in the circuit which in all cases will open before the protector overheats. In most instances this fusible link will be the small conductors in the terminal stub connected to the main cable. In other cases this fusible link will be the drop wire. Thus, the degree of protection to the subscriber has not been changed. Merely a huskier protector with a larger fuse has been provided.

However, the over-all grounding situation is improved. Experience has shown that the fusing of drop wire or small conductors in cable is extremely rare. During fault conditions, therefore, the new protective arrangement of Fig. 2 materially adds to the fault current capacity of telephone plant.

This improvement in electrical protection with the fuseless protector has been accompanied by substantial savings. These savings are twofold. The first cost of the protector installation is lower

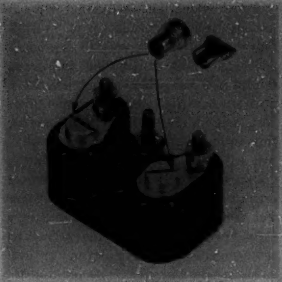


Fig. 3 (left). Fuseless protector for subscriber stations. Detail view is the cylindrical carbon protector block assembly

Fig. 4 (right). Cover for outdoor mounting of the fuseless protector shown in Fig. 3

because the new design permits a reduced material charge for the protector and its housing. After the fuseless protector is in service, a second saving is realized by a reduction in number of maintenance visits to replace operated fuses. This maintenance saving will be particularly apparent in areas of high lightning incidence, since the station fuse performs no useful function in mitigating a lightning surge but is nevertheless extremely vulnerable.

It may be of interest to know how the fuseless protector looks on a customer's premise. Shown in Fig. 3, the unit is not much larger than an egg. The weatherproof housing in Fig. 4 is for outdoor use.

The spark-gap assemblies appear externally as two large screw heads. Whereas this assembly employs air as the dielectric between the carbon electrodes, it is a partially sealed unit and therefore relatively free from moisture troubles inherent in the design of other protector mountings. Another feature of this protector assembly is the low-melting temperature alloy pellet located above the uppermost carbon electrode. With higher currents this alloy softens. The spring base of the mounting can then force both the lower and upper electrode up and make a metallic contact with the brass sleeve of the protector assembly. The arc across the two carbon electrodes is thus short-circuited and extinguished. By this method the amount of heat generated in the protector by an arc is eliminated, and its current capacity becomes a function of the heating of the metallic parts of the protector.

PROVIDING FAVORABLE IMPEDANCES TO POWER CIRCUITS

One of the important considerations in protection is to limit the duration of power voltages on telephone plant and circuits. Theory and experience both indicate that as duration of foreign voltages is reduced, safety to personnel is improved and plant damage is lessened. Fault currents must be high with respect to load currents in order to operate power protective devices promptly. Correspondingly, fault current impedances must be low, and the portion of telephone plant through which a fault current might flow must therefore be of low impedance and have adequate current-carrying capacity to facilitate the operation of power circuit de-energizing devices.

Where power circuits with effectively grounded neutrals exist, bonding the sheaths of telephone cables to the neutral at regular intervals provides a very low impedance from phase wire to neutral in the event of contact between power circuits and telephone plant. In such cases, a phase-to-cable fault is practically the same as a phase-to-neutral fault. For aerial wire the low-impedance path is provided by connecting the heavy duty protectors described earlier to the grounded neutral. With these methods, most of the fault current returns to the power system through the neutral wire rather than through the ground path.

Where the power circuit does not have an effectively grounded neutral, other paths to ground must be relied upon to de-energize the power circuit in the event of a contact. In this case, the impedance presented to the power circuit is the combined impedance of the many paths to ground over the telephone sheath and telephone circuits. Therefore de-energization is frequently less rapid, and telephone plant damage, in general, will be greater. Recalling the discussion on the fuseless protector, this is an instance where the improved current-carrying capacity of the new protector would improve such an undesirable situation. If the trend toward multigrounded neutral power systems continues in the future as it has in the past, these less favorable situations should be reduced in number and degree.

POWER CIRCUIT DE-ENERGIZATION

As indicated in the foregoing, if bonding and grounding are carried out properly, a contact between a power circuit and a telephone circuit produces essentially a short circuit from phase to neutral. Such a short circuit will aid in assuring that the power circuit will be de-energized in a reasonably short time.

The actual de-energization of a power circuit requires, of course, that there be some automatic de-energizing means included in the power circuit. These de-energizing means may take various forms—the important point is that they reliably and promptly interrupt power flow in the event of a low-impedance fault.

It is a characteristic of many power distribution circuits that their load is largely single-phase and hence the so-called 3-phase distribution circuit can be looked upon as 3 single-phase circuits (usually with a single neutral) carried on the same pole head. In most instances a contact between power and telephone plant will involve only one phase wire, and the de-energization means must be arranged to operate under this condition. Furthermore, de-energization requires that the power circuit fuses or relays in each phase be able to distinguish between a normal load and a fault. The ability to do this generally depends on the ratios of currents in the two cases. For usual designs or power circuits the ratio becomes more favorable as normal voltages are increased.

Up until recent years an adequate ratio between short circuit and normal load currents on power systems, at least in cities, was possible at relatively low distribution voltages in the order of 4-kv phase to phase. As power loads have grown and normal load currents have been increased, the ratio has become less favorable. For a given power rating, a higher voltage power system will have a more favorable current ratio, a fact which is borne out by our experience with lines that are joint with 12-kv power.

In many rural areas the time has long since passed when fully satisfactory ratios could be obtained with 4-kv circuits. It is indeed fortunate, from the protection standpoint, that in the relatively early days the power industry adopted voltages in the order of 12 kv for most rural distribution. Experience quite clearly shows that prompt or even reliable de-energization of low-voltage rural distribution circuits is, in many cases, very difficult to obtain.

Effects of New Types of Telephone Wire Insulation

The effect which some of the new materials such as polyethylene will have when used as insulation on individual telephone wires may be of interest. As the relative cost of these materials becomes lower, the amount of their usage will multiply.

In cables these plastic insulating materials compare favorably with paper insulation and provide an advantage in insulation breakdown strength by a factor of from 5 to 20. When very short disturbances, such as lightning, are involved the damage to polyethylene insulated wire will be either nonexistent or will be limited to one or two pairs at the most. On the other hand, when the energy and time of a disturbance are great enough to melt the insulation, there will be little difference between paper and plastic as to the amount of damage done and the number of conductors involved. In the design of telephone cable plant, therefore, the need to provide an outer covering which is well grounded and bonded to the electric neutral remains unchanged.

With the increase in dielectric strength provided by polyethylene insulation many situations where it would otherwise be difficult to protect cable against lightning damage, can be adequately protected. In other more favorable situations where present methods of cable protection are adequate the higher dielectric strength will eliminate the requirements for the protectors that are frequently connected between cable conductors. It will also simplify the design of the protectors by permitting the use of a spark gap in the protector with a higher breakdown voltage and consequently with a larger and less critical gap spacing.

Polyethylene and polyvinyl chloride are also used for insulation on multiple pair distribution wire which has no outer covering. This distribution wire is intended for use where short lengths of open wire or small size cable would otherwise be used. With this type of construction an insulated plant is provided which will withstand most power distribution voltages currently employed. However, present forms of distribution wire are subject to abrasion and possible dielectric breakdown. Therefore, the protective arrangement should not depend entirely on the insulation. It is expected that the probability of a power contact at a point where the insulation is abraided will be much reduced as compared with the similar situation involving uninsulated wires.

◆

No Discussion

Servicing Center for Short-Haul Carrier Telephone Systems

A. L. BONNER
NONMEMBER AIEE

ONE of the principal reasons for designing the various equipment units of *N*,[1] *O*,[2] and *ON*[3] carrier telephone systems on a plug-in basis was to facilitate normal maintenance of these systems. Thus, a faulty unit can easily be replaced by a spare to restore service quickly. The unit in trouble may then be repaired later. The more serious repair work may be done at centrally located "servicing centers," where it can be done efficiently. Since each individual unit performs at least one complete function, it can be completely tested separately.

As will be recalled, the *N* system is a 4-wire 12-channel system for short-haul cables. Because each channel transmits both side bands and the carrier in an 8-kc bandwidth, it can use very inexpensive filters.

The *O* carrier is a 2-wire system for relatively short lengths of open wire. There may be 4, 8, 12, or 16 channels. Because a type-*O* channel utilizes a 4-kc bandwidth to transmit a single side band, the channel filters are relatively expensive compared to *N* carrier filters.

The *ON* carrier utilizes both *N* and *O* equipment to provide 20 channels for short 4-wire cable circuits. Each channel utilizes a 4-kc bandwidth to transmit a single side band. This is accomplished by the use of slightly modified *O* carrier terminal equipment. An *ON* terminal also includes *ON* repeaters (or translators) to translate the terminal frequency spectrum, so that standard *N* repeaters can be used on the line. The equipment may also be arranged for combination open-wire and cable circuits and to connect either one or two 20-channel *ON* systems or terminals to a microwave radio system.

Since where equipment is concerned, these three carrier systems have many features in common, it is desirable that the servicing center be designed to test the equipment of all three systems.

Fig. 1. Some of the carrier units to be tested

Paper 54-288, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 18, 1954; made available for printing May 6, 1954.

A. L. Bonner is with Bell Telephone Laboratories, Inc., Murray Hill, N. J.

Fig. 2. An arrangement for a complete servicing center

Fig. 3. Front view of the test set

Equipment to Be Tested

The principal *N* carrier equipment which can be tested at the servicing center consists of two types of message channel units, high- and low-frequency group transmitting units, high and low group receiving units, and high-low and low-high repeaters. Each of the message channel units has three subassemblies which are tested separately. Other *N* carrier equipment to be tested consists of three types of program channel units, a through channel unit, deviation regulator units, and group and repeater switching sets.

The *O* carrier equipment to be tested consists of channel units (each having three subassemblies), twin channel units, group and repeater oscillators, two types of transmitting group circuits, two types of receiving group and repeater circuits, and several types of filters.

ON equipment consists of *O* terminal equipment, *N* repeaters, and *ON* repeaters.

Altogether there are about 40 different units to be tested. Since some of the units perform several functions, about 90 separate test setups are required. In addition, there are about 20 type-*O* carrier filters to be tested. A group picture of some of the units to be tested is given in Fig. 1.

Requirements for Test Center

A principal requirement for the servicing center is that it should be capable of making essentially the same tests as are made in the factory. It should be sufficiently automatic to set up the tests quickly, yet low enough in price.

In the factory, modulation tests are made on all of the *N*, *O*, and *ON* repeater and group units to ensure that amplifiers have the required feedback characteristics. Since modulation tests require many different filters because of the various frequencies involved, it was decided to substitute measurements of gain around the feedback loops of the amplifiers.

Description of Complete Servicing Center

Fig. 2 shows an arrangement of a complete servicing center. The principal test equipment consists of a servicing center test set, oscillator, vacuum-tube voltmeter, frequency counter, oscilloscope, and volt-ohm milliammeter. An *O* carrier twin-channel unit is shown, held by an adjustable clamp and connected to the test set by a test cord. Any of the test conditions for the unit can be obtained by turning the switch dials involved to their proper position. If trouble is indicated, it can usually be localized by measuring voltages at certain strategic points in the unit. For such work the "send" or input power is continuously applied, and the vacuum-tube voltmeter equipped with test picks is used to measure the voltages. Charts which show exactly where to locate the

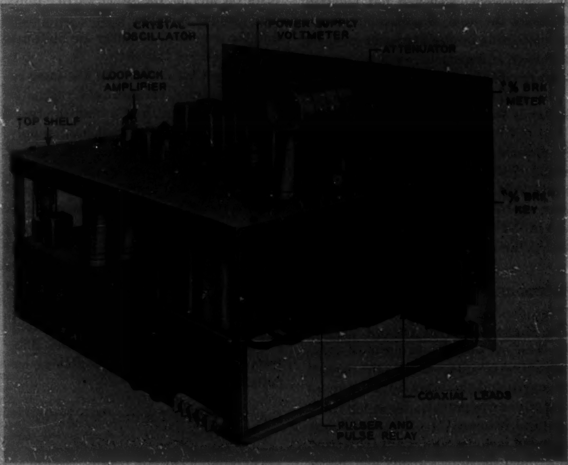


Fig. 4. Side view of the test set chassis

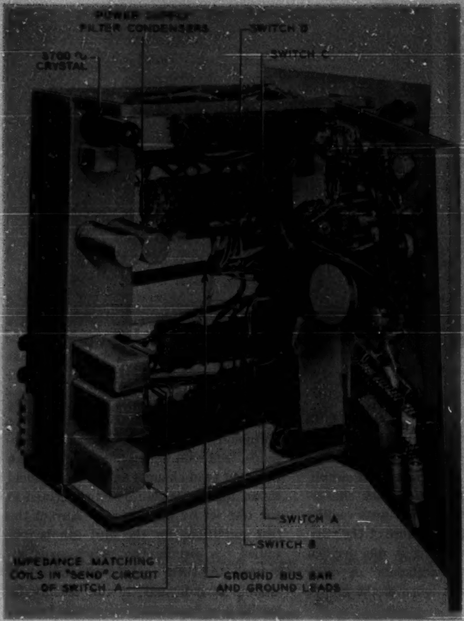

POWER SUPPLY
FILTER CONDENSERS
SWITCH D
3700 ~
CRYSTAL
SWITCH C
SWITCH A
SWITCH B
IMPEDANCE MATCHING
COILS IN "SEND" CIRCUIT
OF SWITCH A
GROUND BUS BAR
AND GROUND LEADS

Fig. 5 (above). Test set chassis with the top shelf swung open

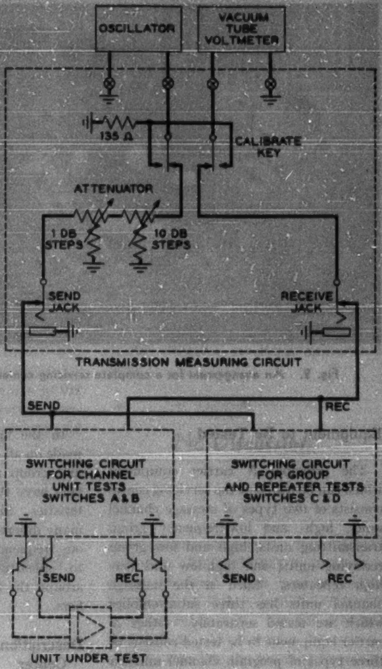

OSCILLATOR
VACUUM
TUBE
VOLTMETER
135 Ω
CALIBRATE
KEY
ATTENUATOR
1 DB
STEPS
10 DB
STEPS
SEND
JACK
RECEIVE
JACK
TRANSMISSION MEASURING CIRCUIT
SEND
REC
SWITCHING CIRCUIT
FOR CHANNEL
UNIT TESTS
SWITCHES A & B
SWITCHING CIRCUIT
FOR GROUP
AND REPEATER TESTS
SWITCHES C & D
SEND
REC
SEND
REC
UNIT UNDER TEST

Fig. 6. Transmission-measuring circuit, and method of making transmission tests

strategic points, where these points are in the circuit, and what the voltages should be, greatly facilitate such tests. The volt-ohm milliammeter is a valuable aid in locating almost all kinds of trouble. The oscilloscope, although not required for any of the tests, is particularly valuable in locating trouble in circuits in which wave form is important, such as the signaling circuit.

The principal use of the frequency counter is to check the frequency of the *O* and *ON* carrier supply oscillators. It will probably be provided wherever the amount of work on these oscillators justifies the cost. If it is not provided, the frequency of the signal oscillators can be checked by a 3,700-cycle crystal filter which is furnished on an optional basis.

Description of the Test Set

Fig. 3 shows a front view of the test set. This set includes two almost independent test circuits: one tests the channel equipment, and the other tests the group and repeater equipment. Such a division is natural because the channel unit test equipment includes equipment for testing the channel signaling circuits and loop-back amplifier, which are not required tests on any of the other carrier units. The channel signaling testing equipment and the loop-back amplifier are built into the test set.

Both test set circuits use the same external oscillator, vacuum-tube voltmeter, frequency counter, and a built-in crystal oscillator for carrier frequency supply. Each test set includes a switching circuit which connects d-c power, carrier supply, signal circuit test equipment, and transmission-measuring equipment as required.

Figs. 4 and 5 show the chassis removed from its cabinet and designate the general location of most of the component parts of the set. The top shelf can be raised to make all of the wiring accessible.

The following describes the main component parts of the test set and illustrates the functions of the switching circuits.

Equipment for Testing Signaling Circuits

The equipment for testing the *N*, *O*, and *ON* carrier channel signaling circuits consists of an on hook-off hook key circuit, a continuous source of dial pulses, and a per-cent break meter circuit to measure how faithfully the channel unit signaling circuit reponds to the dial pulses.

The pulse source is a multivibrator which drives a mercury relay. The dial pulses are obtained from the contacts of the mercury relay and have a fixed rate of 12 pulses per second and fixed per-cent break of 58 per cent. A calibration of the per-cent break circuit is provided.

The per-cent break meter circuit employs a 1-milliampere full-scale meter which is so damped that, when a series of uniform 1-milliampere pulses is applied, the response of the meter is proportional to the percent duration of the pulses. The meterscale is calibrated to give direct readings of per-cent break.

Loop-Back Amplifier

The loop-back amplifier is used to connect (loop back) the output of the channel modulator to the input of the channel

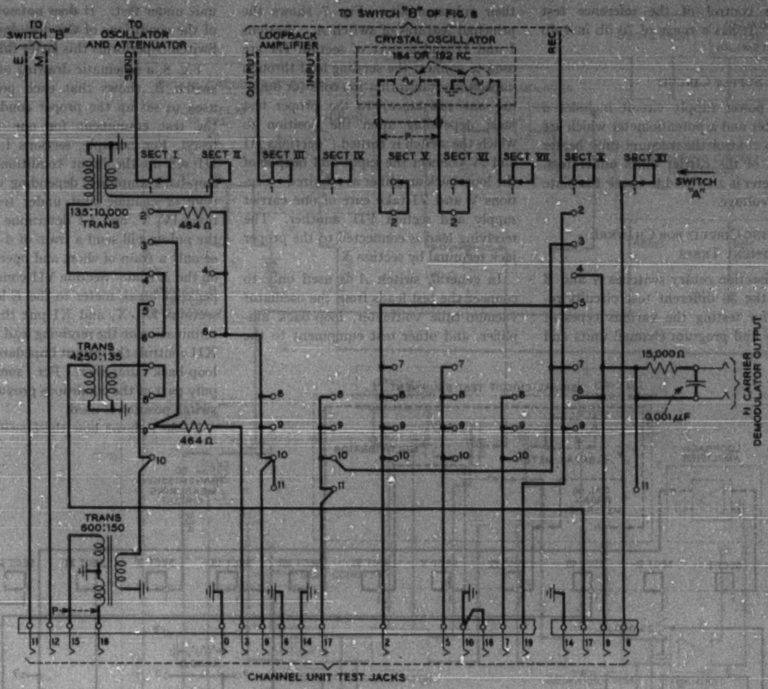

TO SWITCH "B"
E
M
TO OSCILLATOR AND ATTENUATOR
SEND
TO SWITCH "B" OF FIG. 6
LOOPBACK AMPLIFIER
OUTPUT
INPUT
CRYSTAL OSCILLATOR
184 OR 192 KC
REC
SECT I
SECT II
SECT III
SECT IV
SECT V
SECT VI
SECT VII
SECT X
SECT XI
SWITCH "A"
P
135:10,000 TRANS
464 Ω
TRANS 4250:135
464 Ω
TRANS 600:150
P
15,000 Ω
0.001 μF
≈ CARRIER DEMODULATOR OUTPUT
CHANNEL UNIT TEST JACKS

Fig. 7. Principal features of switch A

Table I. Switch Settings for Channel Unit and Channel Unit Subassembly Tests

Carrier Type	Test Setup	Switch A	Switch B	XTAL Oscillator Switch
O and *N*	Receiver potential loss	2	3	
O	Compressor gain	3	2	
N	Compressor gain	3	8	
O and *N*	Expandor gain	4	3	
O and *N*	Keyer loss	5	3	
O and *N*	Signaling receiver gain	6	3	
O and *N*	Signaling receiver adjustment, per-cent break	6	4	
O	Modulator loss	7	5	184 kc
N (*FA*)	Modulator loss	7	9	
N (*F*)	Modulator loss	7	10	
O	Modulator and demodulator loss	8	6	184
N (*FA*)	Modulator and demodulator loss, step (a)	11	11	
	step (b)	11	9	
N (*F*)	Modulator and demodulator loss, step (a)	11	12	
	step (b)	11	10	
O	Over-all channel gain	9	3	184
N (*FA*)	Over-all channel gain, step (a)	9	11	
	step (b)	9	13	
N (*F*)	Over-all channel gain, step (a)	9	12	
	step (b)	9	14	
O	Over-all channel signal-receiving gain	10	6	184
N (*FA*)	Over-all channel signal-receiving gain, step (a)	10	11	
	step (b)	10	9	
N (*F*)	Over-all channel signal-receiving gain, step (a)	10	12	
	step (b)	10	10	
O	Over-all channel signal-receiving adjustment, per-cent break	10	7	184
N (*FA*)	Over-all channel signal-receiving adjustment, per-cent break	10	13	
N (*F*)	Over-all channel signal-receiving adjustment, per-cent break	10	14	

ON channel tests are the same as for *O*. There are two types of *N* units: *F* and *FA*. Whenever the loop-back amplifier is used with *N* channel units, two steps are required: step (a) to adjust the amplifier gain and step (b) to make the measurement.

demodulator. Three different input impedances and two output impedances are required. A special filter is provided for *O* carrier channel units so that the same frequency band can be used for the modulator and demodulator and thereby make loop-back measurements possible.

Crystal Oscillator

The crystal oscillator provides the proper carrier supply frequency for the modulator (or demodulator) of certain *O* carrier units under test. Any one of four different crystals can be switched in to give frequencies of 184, 192, 198, or 236 236 kc.

Transmission Measuring Circuit

Fig. 6 shows a schematic drawing of the transmission-measuring circuit and of the method for making transmission measurements. The calibrate key is used to connect the external vacuum-tube voltmeter to the external oscillator output so that the oscillator output control dial can be set to give a reference test power—usually 0 decibel (db) referred to 1 milliwatt. The variable attenuator permits definite control of the reference test power. It has a range of 70 db in 1-db steps.

Power Supply Circuit

The power supply circuit includes a voltmeter and a potentiometer which are used to control the vacuum-tube heater current of the carrier units under test. This meter is also used to check the plate supply voltage.

Switching Circuit for Channel Equipment Tests

Multisection rotary switches *A* and *B* set up the 36 different test circuits required for testing the various types of message and program channel units and their subassemblies. Fig. 7 shows the principal features of switch *A*. Sections I and II and a part of section III are used to connect the sending lead through impedance transformation coils (or building out resistance) to the proper test jack, depending upon the position to which the switch is turned. Sections III and IV connect the input and output of the loop-back amplifier as required. Sections V and VI take care of one carrier supply and section VII, another. The receiving lead is connected to the proper jack terminal by section X.

In general, switch *A* is used only to connect the test leads from the oscillator vacuum-tube voltmeter, loop-back amplifier, and other test equipment to the unit under test. It does not set up any of the conditions of this test equipment. Switch *B* performs this latter function.

Fig. 8, a schematic drawing of most of switch *B*, shows that each position is used to set up the proper conditions of the test equipment for one or more tests. For example, sections I, II, and III set up the input conditions of the loop-back amplifier, depending upon the type of channel unit under test. Sections IV, V, and VI determine whether the pulser will send a train of d-c pulses, or only a train of short and open circuits on the *M* lead. Section VII connects the per-cent break meter to the *E* lead, and sections IX, X, and XI put the proper termination on the receiving lead. Section XII controls the output impedance of the loop-back amplifier. For some tests only part of the conditions provided by a switch position is used.

Table I shows how these switches are

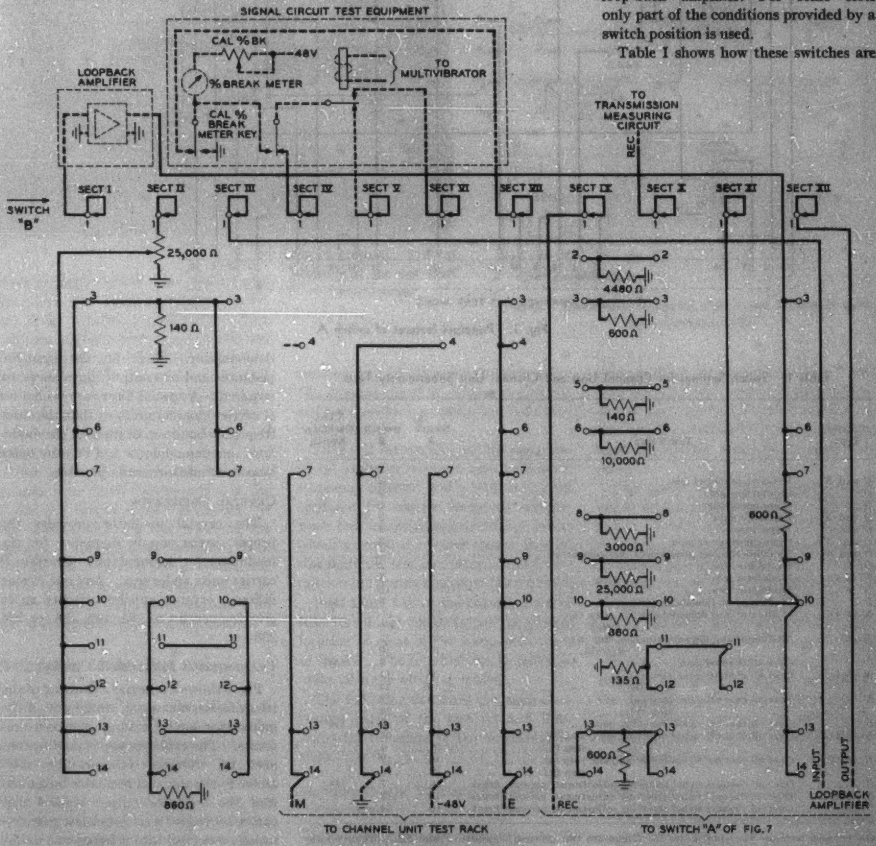

SIGNAL CIRCUIT TEST EQUIPMENT
CAL % BK
-48V
TO MULTIVIBRATOR
% BREAK METER
LOOPBACK AMPLIFIER
CAL % BREAK METER KEY
TO TRANSMISSION MEASURING CIRCUIT
REC
SWITCH "B"
SECT I
SECT II
SECT III
SECT IV
SECT V
SECT VI
SECT VII
SECT IX
SECT X
SECT XI
SECT XII
25,000 Ω
140 Ω
4480 Ω
600 Ω
140 Ω
10,000 Ω
3000 Ω
25,000 Ω
860 Ω
135 Ω
600 Ω
600 Ω
860 Ω
M
-48V
E
REC
INPUT
OUTPUT
LOOPBACK AMPLIFIER
TO CHANNEL UNIT TEST RACK
TO SWITCH "A" OF FIG. 7

Fig. 8. Principal features of switch B

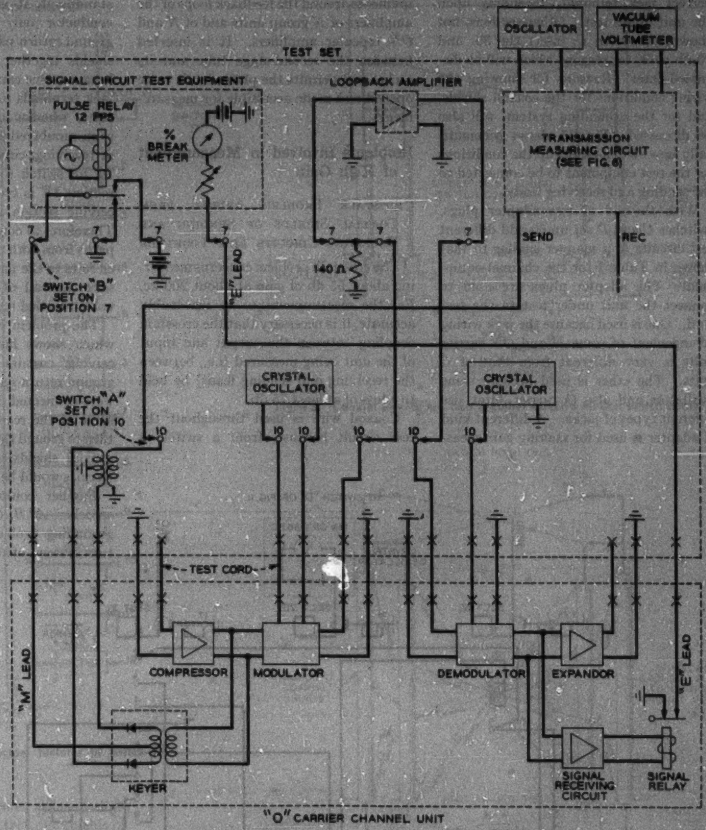


Fig. 9. Tests on type-O carrier channel unit signaling circuit

OSCILLATOR
VACUUM TUBE VOLTMETER
TEST SET
SIGNAL CIRCUIT TEST EQUIPMENT
PULSE RELAY 12 PPS
% BREAK METER
LOOPBACK AMPLIFIER
TRANSMISSION MEASURING CIRCUIT (SEE FIG.6)
7
SEND
REC
140 Ω
"E" LEAD
SWITCH "B" SET ON POSITION 7
CRYSTAL OSCILLATOR
SWITCH "A" SET ON POSITION 10
10
TEST CORD
COMPRESSOR
MODULATOR
DEMODULATOR
EXPANDOR
"M" LEAD
KEYER
SIGNAL RECEIVING CIRCUIT
SIGNAL RELAY
"O" CARRIER CHANNEL UNIT

set to give the correct test conditions for the various message channel units and their subassemblies. It will be noted that the switching circuit is simplified and made more flexible by making each switch perform a single function. For instance, transmission tests on the *O* and both types of *N* channel units (*F* or *FA*) require position 10 only on switch *A*. The different conditions required in the test equipment are met by switch *B*.

Fig. 9 shows the method by which switches *A* and *B* set up test conditions for the signaling circuit of an *O* carrier channel unit. Switch *B*, on step *7*, conditions the signaling test equipment and connects it to the *E* and *M* leads. It also conditions the loop-back amplifier. Switch *A*, on step 10, connects the 3,700-cycle signaling supply to the keyer in the channel unit and also connects the carrier frequency supplies and loop-back amplifier to the channel unit.

With the switches set as described, the test is under way. The 3,700 cycles into the keyer will be pulsed by the alternate action of the battery and ground on the *M* lead, stopping and restoring transmission through the varistors in the keyer. Actual dial pulsing is thus simulated. The pulses' 3,700 cycles are passed through the modulator, loop-back amplifier, and demodulator into the signaling receiving circuit where they operate the signal-receiving relay. The per-cent break meter circuit, now connected to the *O* carrier signal relay contacts by the *E* lead, is used to determine whether the pulsing relay faithfully follows the pulse applied at the *M* lead.

Switching Circuit for Group, Repeater, Oscillator, and Filter Tests

Switches *C* and *D* set up the conditions required for testing *O*, *ON*, and *N* carrier group repeater, oscillator, and filter equipment previously described. In general, the tests consist of gain (or loss) measurements. As in the case of the channel unit tests, one switch is used for making connections to the unit under test and the other switch is used for setting up the various conditions required for the test equipment.

Fig. 10 shows actual wiring conditions for some of the transmission tests performed by switch *C*. It shows how the sending and receiving leads and the carrier supply leads are connected to the correct jack terminal, depending upon the unit under test. Other sections, not shown, of switch *C*, insert the 30- and 20-db pads, for reasons which will be discussed later. Reasons for showing the wiring conditions for the coaxial shields, and for the grounding system, will also be discussed. Fig. 11 shows schematically how switch *D* sets up the conditions for the test equipment to be connected to the sending and receiving leads.

With the aid of two adapter plugs, switches *C* and *D* set up 60-odd different test circuits in a manner similar to that shown in Table I for the channel equipment. The adapter plugs are used to connect the unit under test to the test cord. One is used because the jack wiring arrangement of some *N* and *ON* carrier units is very different from that of *O* units. The other is used because some oscillators and also *O* carrier filter use different types of jacks. A different kind of adapter is used for making gain measurements around the feedback loops of the amplifiers of *N* group units and of *N* and *ON* repeater amplifiers. It is inserted between the second-stage tube and its socket and permits the plate circuit to be opened and made accessible for measurements.

Problems Involved in Measurement of High Gain

Crosstalk Problem Arising from Coaxial Shields of Sending and Receiving Circuits Tied Together

The crosstalk problem concerns measuring about 55 db of gain at about 200 kc. For the measurement to be reasonably accurate, it is necessary that the crosstalk coupling between the output and input of the unit being measured (i.e., between the receiving and sending leads) be held to a loss of at least 75 db.

Coaxial wire is used throughout the test circuit because, from a switching standpoint, it is simpler to switch the conductor only and to complete the ground return paths by tying the coaxial shields together, a complicated process because some carrier units employ certain jack terminals for input, or sending, circuits, whereas other units employ the same numbered jack terminals for output, or receiving, circuits. For example, Fig. 10, for switch *C*, shows how the sending section *IV* is cross-connected to the receiving section *XI*, and other sections. Therefore, to complete the ground return paths from both the sending and receiving circuits to the particular jack, it is necessary that all of the coaxial shields involved be tied together in some manner.

The problem is illustrated by Fig. 12, which shows how the sending and receiving circuits would have common ground return paths if the coaxial shields were connected directly together. Since some of the receiving current would return to ground through the sending circuit coaxial shields, the loss between these circuits would be very low.

Further complications arise because switches, *A*, *B*, *C*, and *D* have a similar grounding problem and the ground return systems of all four switches are inter-

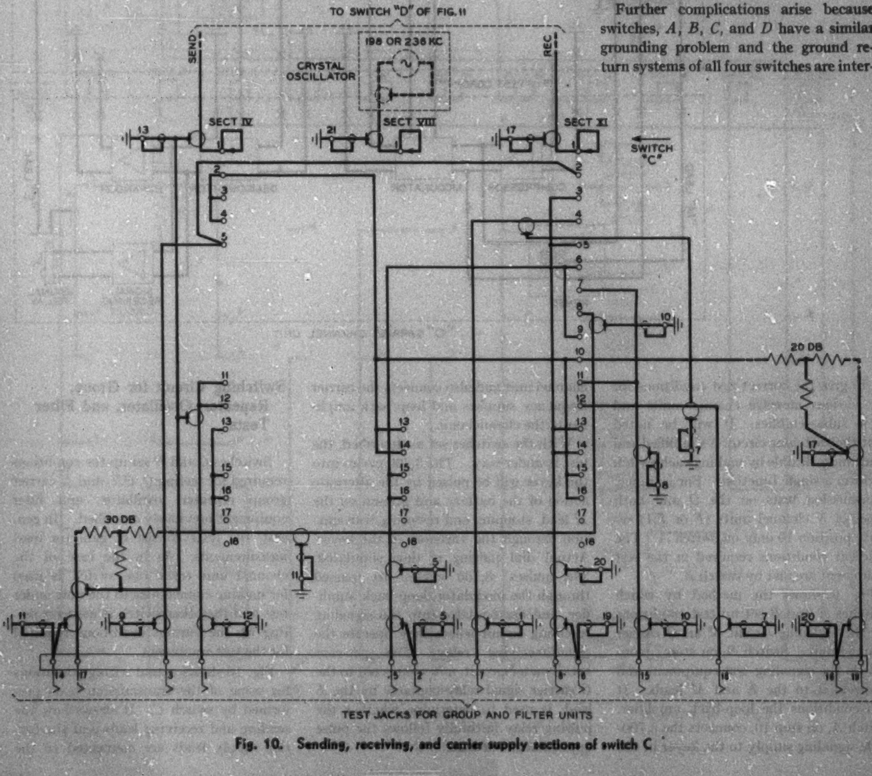


Fig. 10. Sending, receiving, and carrier supply sections of switch C

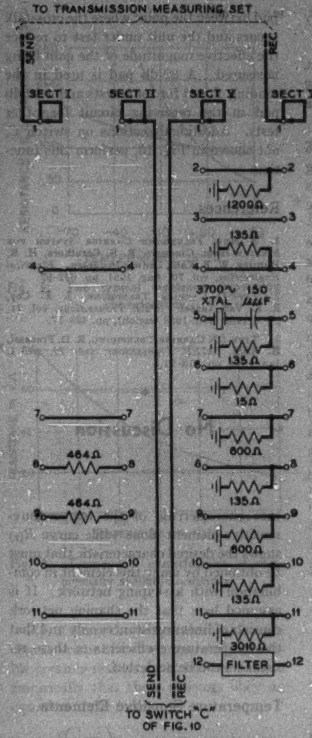

TO TRANSMISSION MEASURING SET
SEND
REC
SECT I
SECT II
SECT V
SECT VI
SWITCH "D"
1
2
3
4
5
6
7
8
9
10
11
12
1200Ω
135Ω
3700~ XTAL
150 μμF
135Ω
15Ω
600Ω
135Ω
600Ω
135Ω
3010Ω
464Ω
464Ω
FILTER
SEND
REC
TO SWITCH "C" OF FIG. 10

Fig. 11. Principal features of switch D

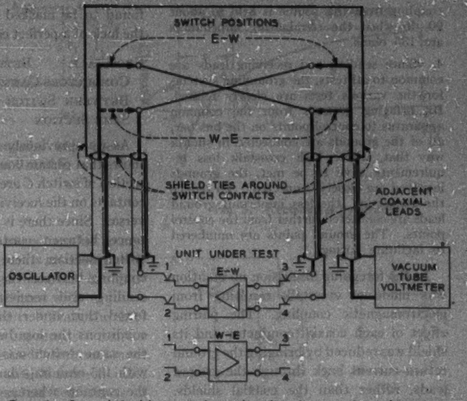

SWITCH POSITIONS
E-W
W-E
SHIELD TIES AROUND SWITCH CONTACTS
ADJACENT COAXIAL LEADS
UNIT UNDER TEST
OSCILLATOR
VACUUM TUBE VOLTMETER
E-W
W-E
1
2
3
4

Fig. 12. Coaxial shields are tied together directly around switch contacts

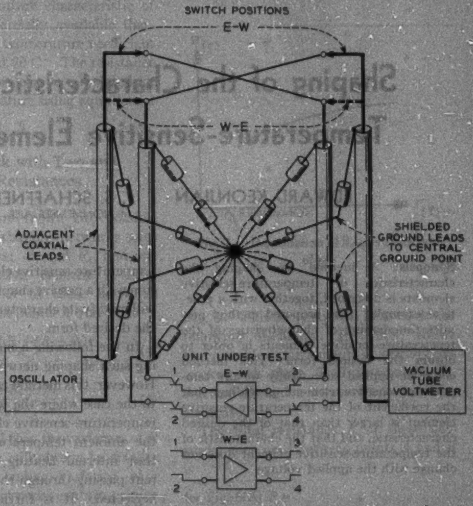

SWITCH POSITIONS
E-W
W-E
ADJACENT COAXIAL LEADS
SHIELDED GROUND LEADS TO CENTRAL GROUND POINT
UNIT UNDER TEST
OSCILLATOR
VACUUM TUBE VOLTMETER
E-W
W-E
1
2
3
4

Fig. 13. Coaxial shield return paths are completed through common ground point

connected by a common transmission-measuring circuit. In fact, the shields of almost all the coaxial leads must be tied together in some way.

The problem of common impedance was solved by tying each end of each coaxial shields to a common ground point. The ground leads used to do this have grounded shields. Fig. 10, for switch *C*, gives an example of actual wiring conditions. The principles involved are illustrated by Fig. 13 and are discussed in the following:

1. The principal source of crosstalk between the coaxial leads is caused by the capacitances between them. In Fig. 13 all of these coupling capacitances are short-circuited by the ground leads to the common ground point. It was found considerably more effective to tie both ends of the coaxial shields to ground, rather than just one end, because the coaxial shields have an appreciable resistance. Grounding both ends of the shields more nearly short-circuits the capacitances.

2. Fig. 13 shows that the ground leads form the return path for much of the return current. At the switch points they form the only path. Since these leads lie close together at the ground bus bar and in some of the cabling, it was found necessary to shield them, from a crosstalk standpoint. One end of each shield is grounded at the ground bus bar at the same point as the ground lead. The grounded shield reduces capacitive coupling between the ground leads, and the spacing between leads provided by the use of shielded wire reduces electromagnetic coupling.

3. The principal source of crosstalk between the sending and receiving circuits is attributable to the impedance of the common ground point provided by the bus bar. A section of bar about 2½ inches long and 5/8 inch wide was used. It has 28 holes about 3/16 inch apart for connecting the 50-odd ground leads. A picture of the ground leads and bus bar is shown in Fig. 5. In tests where rigid crosstalk requirements have to be met, the ground leads involved are brought together in a bus bar area about 3/4 inches in diameter. Thus, the 200-kc crosstalk

coupling from this source is held to about 90 db when the terminating impedances are 135 ohms.

4. Since sending and receiving leads are common to all tests, the grounding systems for the various tests are closely related. By bringing grounds from the common apparatus to control points on the bus bar, all of the grounds are connected in such a way that, where high crosstalk loss requirements have to be met, the grounds involved are closely connected. To reduce the congestion, the less important ground leads are connected further from the control points. The ground points are numbered to facilitate wiring.

In the foregoing discussion no mention was made of crosstalk resulting from electromagnetic coupling. The pairing effect of each coaxial conductor and its shield was reduced by bringing the ground return current back through the ground leads, rather than the coaxial shields, thereby increasing the electromagnetic crosstalk. However, this crosstalk was found to be masked by that caused by the lack of a perfect common ground.

CROSSTALK BETWEEN COAXIAL CONDUCTORS CAUSED BY CAPACITANCE BETWEEN SWITCH CONTACTS ON THE SAME SECTION

As was previously mentioned, Fig. 10 shows that certain contacts of the sending section of switch *C* are also tied to various contacts on the receiving section, and vice versa. Since there is appreciable capacitance between contacts of the same switch section, there is some crosstalk coupling between the conductors of the sending and receiving leads. It was found that under the worst humidity conditions the loss between contacts on the same switch was 55 db at 200 kc, with 135-ohm impedance terminations at the contacts whereas the over-all crosstalk requirement is 75 db.

This problem was solved by putting a pad between the point where the crosstalk occurs and the unit under test to reduce the effective magnitude of the gain being measured. A 30-db pad is used in the sending circuit for some tests and a 20-db pad in the receiving circuit for other tests. Additional sections on switch *C*, not shown in Fig. 10, perform this function.

References

1. A NEW TELEPHONE CARRIER SYSTEM FOR MEDIUM-HAUL CIRCUITS, R. S. Caruthers, H. R. Huntley, W. E. Kahl, Ludwig Pedersen. *Electrical Engineering*, vol. 70, Aug. 1951, pp. 692–97.

2. TYPE-O CARRIER TELEPHONE, J. A. Coy, E. K. Van Tassel. *AIEE Transactions*, vol. 71, pt. I, 1952 (Jan. 1953 section), pp. 428–37.

3. TYPE-ON CARRIER TELEPHONE, R. D. Fracassi, H. Kahl. *AIEE Transactions*, vol. 72, pt. I, Jan. 1954, pp. 713–21.

◆

No Discussion

Shaping of the Characteristics of Temperature-Sensitive Elements

EDWARD KEONJIAN
NONMEMBER AIEE

J. S. SCHAFFNER
NONMEMBER AIEE

Synopsis: A method for shaping the characteristics of temperature-sensitive elements is described, together with a practical example. The proposed method permits the shaping of characteristics of the temperature-sensitive elements in order to obtain the desired form. This method could be applied successfully to any temperature-sensitive element providing that the coefficient of the temperature-sensitive element is larger than that of the desired characteristic, and that the characteristic of the temperature-sensitive element does not change with the applied voltage.

IN RECENT years, temperature-sensitive elements have found many useful applications, particularly in the field of control devices. In general, these elements are used in two ways: 1. in electrical systems, controlled by changes in temperature; 2. as means for the compensation of undesirable effects caused in electrical networks by temperature variations.

In general, the characteristics of the commercially available temperature-sensitive elements do not have the form that is required for a specific application. Therefore it may be necessary to use the temperature-sensitive element in combination with a passive shaping network which will modify its characteristic so that it has the desired form.

In the following a method for calculating such shaping networks will be given. However, this method can be applied only to the case where the temperature of the temperature-sensitive element is equal to the ambient temperature. This implies that internal heating (caused by current passing through the element) can be neglected. It is further assumed that for a given ambient temperature the resistance of the elements is constant and therefore independent of the current passing through them.

In Fig. 1, curve $r_{(T)}$ shows the temperature characteristic of the temperature-sensitive element alone while curve $R_{(T)}$ shows the desired characteristic that must be obtained by using this element in combination with a shaping network. It is assumed here that the shaping network consists of linear resistances only and that the temperature coefficients of these resistances can be neglected.

Temperature Sensitive Elements

In the realization of the proposed method of temperature compensation, a number of temperature-sensitive elements may be used. Among these elements are special types of ceramic resistors, and thermistors.

CERAMIC RESISTORS

These resistors are made of compositions with marked negative temperature coefficients, approximately 1 to 3 per cent per degree centigrade (C) at 25 C.

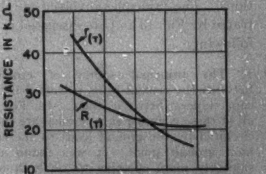


Fig. 1. Temperature characteristics of temperature-sensitive element and desired network

---

Paper 54-269, recommended by the AIEE Electronics Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 23, 1954; made available for printing May 13, 1954.

EDWARD KEONJIAN and J. S. SCHAFFNER are with the General Electric Company, Syracuse, N. Y.

The assistance of the staff of the Electronics Laboratory, General Electric Company, Syracuse N. Y., in particular that of Dr. S. K. Ghandhi, is gratefully acknowledged.

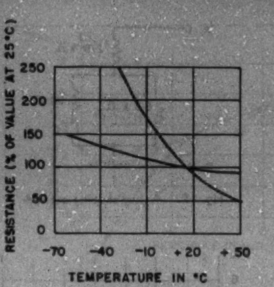


Fig. 2. Two typical resistance-temperature characteristics

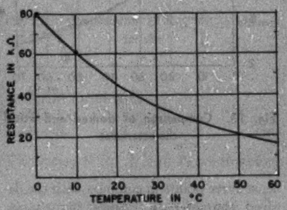


Fig. 3. Temperature characteristics of commercially available resistance

Fig. 2 shows the temperature-resistance characteristic of two commercially available ceramic resistors. It has been found empirically that these resistors obey an equation of the form

$$R=R_0\epsilon^{B\left(\frac{1}{T}-\frac{1}{T_0}\right)} \tag{1}$$

In this equation, T_0 is the reference temperature in degrees Kelvin (for 25 C, $1/T_0=0.003356$), $R=$the resistance at temperature T, $R_0=$the resistance at reference temperature, T equals temperature of the resistor in degrees Kelvin, and B equals the temperature constant. A more convenient form of this equation for computation is

$$\text{Log}\frac{R}{R_0}=0.4343\,^{B}\left(\frac{1}{T}-\frac{1}{T_0}\right) \tag{2}$$

The value of the constant B varies with the resistivity of the composition and can be obtained from the curves supplied by the manufacturer. A typical value of B is a few thousand degrees Kelvin.

The temperature coefficient a of these resistances can be calculated from equation 1

$$a=\frac{1}{R}\frac{dR}{dT}=-\frac{B}{T^2} \tag{3}$$

A typical value for a at room temperature is −3 per cent per C.

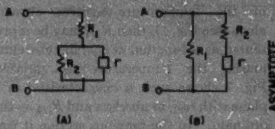


Fig. 4. Shaping networks with two fixed resistances R_1 and R_2

THERMISTORS

The word "thermistors" applies to any resistor which changes its resistance with temperature. Usage, however, has limited the term to resistors with negative temperature coefficients. Thus, the thermistor is a temperature-sensitive resistor consisting of a sintered combination of ceramic materials and various metallic oxides. The temperature coefficients of thermistors is negative and is of the order of 2.5 per cent per C at 20 C. Fig. 3 shows the temperature characteristic of one of the commercially available thermistors, having a temperature coefficient of −2.7 per cent at 20 C. The resistance of these thermistors can be calculated for any given temperature using equations 1 and 2.

Shaping Network with Two and Three Fixed Resistances

In Fig. 4(A) are shown two possible arrangements of shaping networks with two fixed resistors; R_1 and R_2, and one temperature-sensitive element r. The total resistances R of these networks (measured between points A and B) is

$$\text{for Fig. 4(A) } R=\frac{R_1R_2+(R_1+R_2)r}{R_2+r} \tag{4}$$

$$\text{for Fig. 4(B) } R=\frac{R_1R_2+rR_1}{R_1+R_2+r} \tag{5}$$

It should be noted that the variation of R with temperature will always be smaller than the corresponding variation of r alone, since the networks contain passive resistances in series and parallel with the temperature-sensitive element. It is also obvious than an increase in $r_{(T)}$ will always cause an increase in $R_{(T)}$; and a decrease in r, a decrease in $R_{(T)}$. Mathematically this may be expressed as

$$0<dR/dr<1 \tag{6}$$

Equations 4 and 5 are of the general form

$$R_{(T)}=\frac{a+br_{(T)}}{c+r_{(T)}} \tag{7}$$

with positive coefficients a, b, and c. Geometrically this corresponds to rectangular hyperbolas of the form shown in Fig. 5. This property will be used in the following discussion.

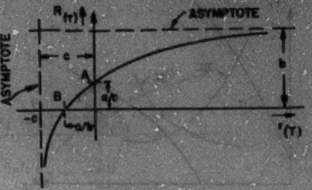


Fig. 5. Rectangular hyperbola

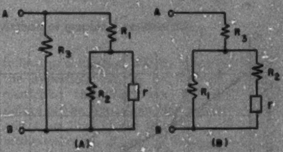


Fig. 6. Shaping networks with three fixed resistances R_1, R_2, and R_3

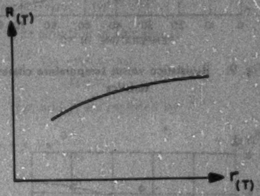


Fig. 7. The variation of $R_{(T)}$ as a function of $r_{(T)}$

Fig. 6(B) shows two possible arrangements of shaping net works with three fixed resistances R_1, R_2, and R_3 and a temperature sensitive element r.

The total resistance R of these networks, measured between points A and B is

$$\text{for Fig. 6(A) } R=R_3\frac{R_1R_2+r(R_1+R_2)}{R_1R_2+R_2R_3+r(R_1+R_2+R_3)} \tag{8}$$

$$\text{for Fig. 6(B) } R=\frac{(R_1R_2+R_1R_3+R_2R_3)+r(R_1+R_3)}{R_1+R_2+r} \tag{9}$$

The inequality, equation 6, holds for equations 8 and 9 and, likewise, their general form is that of equation 7.

Criteria for Possible Shaping Arrangements

If $R_{(T)}$ is the desired temperature characteristic and $r_{(T)}$ the characteristic

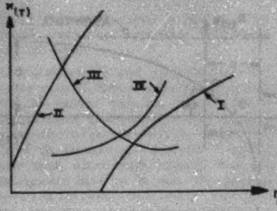


Fig. 8. Undesirable characteristics

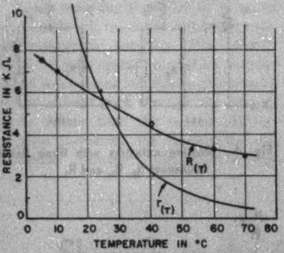


Fig. 9. Resistance versus temperature characteristics

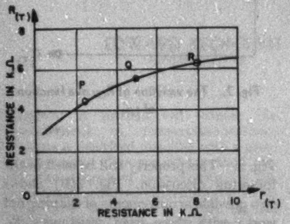


Fig. 10. The $R_{(T)}-r_{(T)}$ characteristic

of the temperature-sensitive element alone (see Fig. 1), then $R_{(T)}$ may be represented as a function of $r_{(T)}$, thereby eliminating the temperature as variable. Fig. 7 shows such a characteristic in a plane with $r_{(T)}$, as abscissa and $R_{(T)}$ as the ordinate. In the previous section it was shown that the relationship between $R_{(T)}$ and $r_{(T)}$ for shaping networks consisting of two or three externally added resistances can be expressed by the equation of a rectangular hyperbola (equation 7).

The asymptotes of this hyperbola are: $R_{(T)}$ equals b and $r_{(T)}$ equals $-c$. As may be seen, this hyperbola intersects the $R_{(T)}$ axis at a point A equals a/c and the $r_{(T)}$ axis at a point B equals a/b. The basic problem of obtaining a desired temperature characteristic $R_{(T)}$ with a given temperature-sensitive element and a number of externally added resistances may now be reduced to that of approximating the curve of Fig. 7 by a rectangular hyperbola as shown in Fig. 5.

A good approximation, however, is not possible in a number of cases. For example, this is true for the characteristic I of Fig. 8 since this characteristic has a negative slope and for the characteristic II in the same figure since the latter is curved upward instead of downward. Curve III has a positive intersect with the $r_{(T)}$ axis and curve IV has a slope larger than unity. In all such cases, where the curve corresponding to a given $T_{(T)}$ and $r_{(T)}$ cannot be approximated readily by a rectangular hyperbola of the type shown in Fig. 5, a new temperature-sensitive element with a different characteristic should be selected.

Synthesis of Temperature-Sensitive Networks

In the following a method will be described for constructing a shaping network with a $R_{(T)}-r_{(T)}$ characteristic that has at least three points in common with the desired characteristic. Since the characteristic of the shaping network is a hyperbola and the desired characteristic may have a considerably different form, it may not generally be expected that the two curves have more than three points in common.

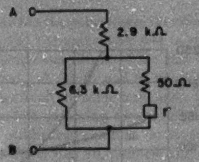


Fig. 12. Actual shaping network

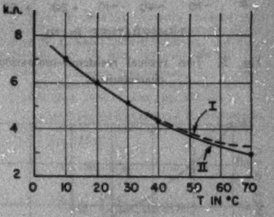


Fig. 13. Comparison of desired and actual characteristics

To approximate a desired temperature characteristic in the $R_{(T)}-r_{(T)}$ plane, three points on this characteristic must be selected. Then a rectangular hyperbola passing through these three points should be constructed. A graphical method for this construction is described in the Appendix. The coefficients a, b, and c of equation 7 describing the hyperbola are determined by the co-ordinates of the origin of the hyperbola and its intersects with the abscissa and ordinate (Fig. 5).

Knowing these coefficients, the elements of the shaping network can be calculated, for example, by comparing equations 7 and 9. These two equations are identical if (see following equation)

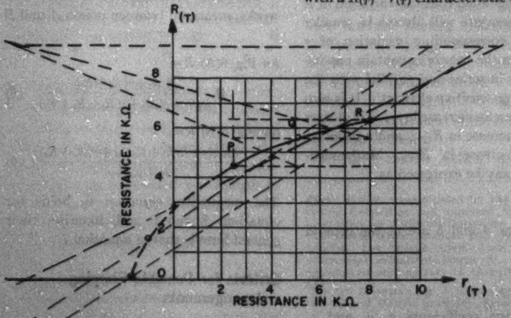


Fig. 11. Construction of the hyperbola

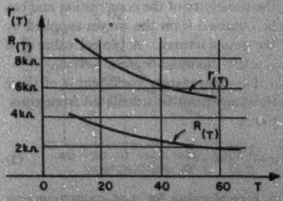


Fig. 14. $R_{(T)}$ and $r_{(T)}$ characteristics

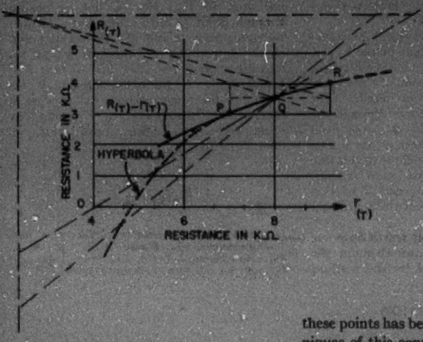


Fig. 15. Construction of $R_{(T)}-r_{(T)}$ characteristic

$$a=R_1R_2+R_1R_3+R_2R_3$$
$$b=R_1+R_3$$
$$c=R_1+R_2 \qquad (10)$$

The values of R_1, R_2, and R_3 are therefore

$$R_1=\sqrt{bc-a}$$
$$R_2=c-R_1$$
$$R_3=b-R_1 \qquad (11)$$

Practical Examples

Fig. 9 shows an example of the two curves, one representing the characteristic of the temperature sensitive element alone, $r_{(T)}$; the other the desired temperature characteristic $R_{(T)}$.

From these two curves, the characteristic in the $R_{(T)}-r_{(T)}$ plane may be obtained as shown in Fig. 10. As shown in the section, "Synthesis of Temperature-Sensitive Networks," this curve should be approximated by a rectangular hyperbola in order to obtain the necessary design information for the shaping network. In Fig. 11 three points of the $R_{(T)}-r_{(T)}$ characteristic have been selected and a rectangular hyperbola passing through these points has been constructed. Techniques of this construction appear in the Appendix. The asymptotes of this hyperbola are $R_{(T)}=9.2k\Omega$ and $r_{(T)}=-6.25k\Omega$, and the intersection with the $R_{(T)}$ axis is $R_{(T)}=2.9k\Omega$. The coefficients a, b, and c of equation 7 are therefore

$$a=18.4(k\Omega)^2;\ b=9.2k\Omega,\ c=6.25k\Omega$$

For a shaping network as shown in Fig. 6(B), the values of the resistances R_1, R_2, and R_3 can be calculated from equation 12.

$$R_1=\sqrt{bc-a}=6.3k\Omega$$

$$R_2=c-R_1=50\Omega$$

$$R_3=b-R_1=2.9k\Omega$$

Fig. 12 shows the complete network. The temperature characteristic of this network is indicated as curve I in Fig. 13. Curve II of the same figure represents the desired characteristic. Comparing the two curves shows that the approximation is close.

A case where $r_{(T)}$ and $R_{(T)}$ are of different orders of magnitude, is shown in Fig. 14.

Fig. 15 shows the $R_{(T)}-r_{(T)}$ characteristic and the hyperbola approximating it.

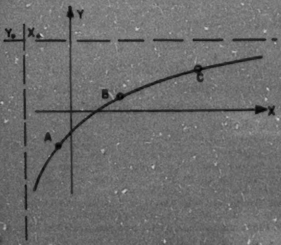


Fig. 16. Rectangular hyperbola

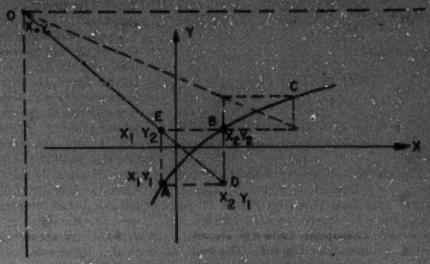


Fig. 17. Method of graphical construction

Appendix. A Method for Constructing a Rectangular Hyperbola from Three Given Points

A rectangular hyperbola with the origin as the center is given by the equation

$$xy=a$$

If the center is at x equals x_0, y equals y_0, this equation is

$$(x-x_0)(y-y_0)=a$$

or

$$y=y_0+\frac{a}{x-x_0}=\frac{y_0x+(a-x_0y_0)}{x-x_0} \qquad (12)$$

Fig. 16 shows such a hyperbola. In the following a method will be given by which the hyperbola can be constructed, and in particular the point x_0y_0 found if three points of the hyperbola (for example A, B, and C in Fig. 1) are known. Let the coordinates of A, B, and C be x_1y_1, x_2y_2 and x_3y_3 respectively.

First it shall be proved that the line connecting the points E and D of Fig. 17 passes through the point x_0, y_0, where E and D have the co-ordinates y_2x_1 and y_1x_2 respectively.

The equation of this line is

$$y=y_1-\frac{(y_2-y_1)(x-x_2)}{x_2-x_1}$$

If y_1 and y_2 are replaced by

$$y_1=y_0+\frac{a}{x_1-x_0}$$

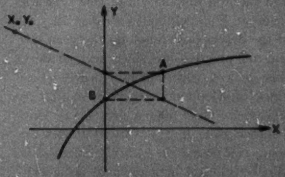


Fig. 18. Construction of intersection with y-axis

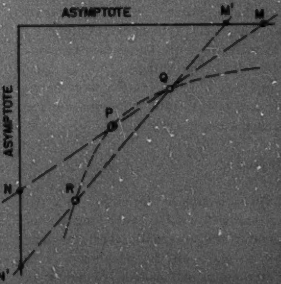


Fig. 19. Construction of hyperbola

$$y_2 = y_0 + \frac{a}{x_2 - x_0}$$

there results

$$y = y_0 + \frac{a(x - x_0)}{(x_1 - x_0)(x_2 - x_0)} \qquad (13)$$

For $x = x_0$, we have $y = y_0$. The dotted line of Fig. 17 therefore passes through the point x_0, y_0.

The method of construction of the point $x = x_0$, $y = y_0$ from three points *A*, *B*, and *C* of the hyperbola based on this is obvious from Fig. 17 where two rectangles are constructed using the known points *A*, *B*, and *C* as corners. The two diagonals of these rectangles must both pass through x_0y_0. The point where they cross is therefore x_0y_0.

In the following the graphical construction of the intersection point of the hyperbola with the *y* axis is given. (See Fig. 18). In this figure, the intersection of the hyperbola with the *y* axis should be found. First a parallel to the *x* axis is drawn through the known point *A*. Then a line is drawn from x_0y_0 through the intersection of the *y* axis and this parallel. A rectangle with this line as a diagonal is then constructed as indicated in Fig. 18. The corner *B* of this triangle is the intersection of the hyperbola with the *y*-axis.

If the asymptotes and at least one point *Q* of a hyperbola, are known then further points can be constructed by the method indicated in Fig. 19. A line is drawn through point *Q* intersecting both asymptotes in points *N* and *M*. A second point of the hyperbola can then be found by locating a point *P* on the line so that *PN* equals *QM*. Additional points may be found by the same methods by drawing a different line through *Q* (or *P*) using the same techniques. (See construction of point *R* in Fig. 19).

Reference

1. Les Proprietes et les Applications des Thermistances, J. Bleuze. *Onde Electrique*, Paris, France, no. 317–318, Aug.-Sept., 1953, pp. 497–500; no. 319, Oct., 1953, pp. 578–90.

No Discussion

Application of Toll Dialing in Pennsylvania

A. L. CHARNY
NONMEMBER AIEE

PLANS for mechanizing the handling of long-distance telephone calls, using common control equipment to obtain improved quality of service and economy of operation, were initiated in 1941 and the first installation of this type was made at Philadelphia in 1943.[1] In spite of the delays occasioned by the war, substantial progress has since been made in this project which affects nearly every phase of telephone operation, so that by 1954 more than 50 per cent of all long-distance calls in the United States are being completed on a mechanized basis. Many technical articles have been written describing the general plan for mechanization, as well as specific parts.[2–7] This paper is intended to cover the problems involved in applying the nation-wide dialing plan to the territory operated by The Bell Telephone Company of Pennsylvania. It should be clear, however, that the independent telephone companies which operate a large part of the area of the State of Pennsylvania are participating in these plans.

Objective

The objective of the nation-wide dialing plan is to enable every long-distance operator and eventually every telephone user to reach all the telephones in the United States and Canada by dialing just as is now done on local calls. The customers will dial their own station-to-station calls, but the operators will continue to handle person-to-person calls, collect calls, calls from coin telephones, and calls involving quotation or transfer of charges.

Dual Problem in Telephone Connections

Each dial telephone connection presents a dual problem. The originating end must be furnished the means of signaling the outward operator or outward automatic equipment; the incoming end must be furnished the means of receiving inward signals and eventually ringing the called telephone. This duality exists in all parts of telephone equipment, the simplest illustration being the telephone set, where the dial is the outward signaling mechanism and the ringer is the incoming signaling device. This fact allows the separation of the problem of mechanization into an outward and an inward component and permits orderly and economical progress in one direction without necessarily affecting the other. At an illustration, the central office at West Chester, Pa., is manual and inward calls are completed manually by an operator, but on outward calls the West Chester toll operators use dials at their toll positions to complete calls directly to stations throughout the United States. The reverse situation may exist where automatic selectors are installed to receive incoming dialed calls, whereas outward calls are still being handled on a ringdown basis without dialing.

The most important application of this principle occurs in connection with the introduction of customer dialing of long-distance calls. At the time the central office is equipped for customer dialing the users are furnished with a list of points which have been equipped to receive incoming dialed calls from customers and they are instructed to dial these points in accordance with the nation-wide plan. There are at present about 21,000,000 stations in the United States equipped for receiving inward customer dialing, but only 42,000 stations are equipped for outward customer dialing, although rapid expansion of this method is contemplated during the coming years.

In Pennsylvania there were, as of January 1, 1954, about 2,440,000 dial telephones, of which 1,847,000 are arranged to receive customer dialing. The number of telephones served by outward operator dialing is about 2,600,000. Of these, 13,000 telephones in the Valley central office in East Pittsburgh are equipped to dial their own long-distance calls.

Prerequisites for Nation-wide Dialing

In addition to the local automatic equipment itself the following arrangements are essential for successful nation-wide dialing:

Paper 54-273, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 29, 1954; made available for printing April 30, 1954.

A. L. Charny is with The Bell Telephone Company of Pennsylvania, Philadelphia, Pa.

1. A nation-wide numbering plan.
2. A system of switching centers.
3. An interconnecting network of circuits.

The first two of these arrangements will be discussed with respect to their application in Pennsylvania.

Nation-wide Numbering Plan

PRESENT CONDITIONS

The existing numbering plans were dictated by the economies of local dialing; as a result they generally vary in accordance with the size of the city, beginning with the familiar 2 letters (L) and 5 numerals (N) in the large cities, 5 numerals or a mixture of 5 and 4 numerals in the intermediate sized cities, and down to 3 numerals in the smallest towns. Other numbering plans use 2 letters and 4 numerals, 1 letter and 5 numerals, etc. Operator toll dialing was planned to fit into the existing telephone numbering arrangements on the assumption that it is more economical to instruct the operators to distinguish among the various telephone station designations than to convert them all to a uniform system. By means of directions printed on a toll switchboard bulletin, one sheet in the case of a small office, or a loose-leaf book for the larger offices, the operators are instructed as to the digits which must be dialed to reach the called telephone. A few entries from the Allentown, Pa., toll switchboard bulletin illustrate the variety of instructions necessitated by the numbering plans at the distant cities.

Destination	Route
Akron, Ohio	Pittsburgh+042+2L+(4) or (5)
Minneapolis, Minn.	Pittsburgh+612+2L+ or 3L+
State College, Pa.	Harrisburg+037+8+
York, Pa.	Philadelphia+143+

To reach a number in Akron, Ohio, the Allentown outward operator plugs into an idle trunk to Pittsburgh. The + sign indicates that the code following it should be dialed. In this case the code 042 will reach an idle trunk from Pittsburgh to Akron and the subsequent dialing of 2L+4N or 2L+5N will complete the call to the Akron number. The Minneapolis entry illustrates a mixture of 2L+4N and 3L+4N in the same city. The entry for State College, Pa., illustrates a call dialed to a tributary office reached via its toll center. The code 037 via Harrisburg reaches Bellefonte, Pa., which is the toll center and the code 8 switches at Bellefonte to State College. To reach York, Pa., the operator selects a trunk to Philadelphia and dials 143 and the 4- or 5-digit number listed in the directory.

The justification for the continuation of this method lies in the need for postponing the large costs of converting the existing central office equipments to the standard numbering system, i.e., 2 letters and 5 numerals. However, even with the training given to toll operators and the instructions available through the switchboard bulletins, the desirability of simplification is apparent. With customer dialing this becomes an absolute necessity. This simplification has been achieved with the nation-wide numbering plan and the use of destination codes.

PROPOSED NUMBERING PLAN

Under the nation-wide numbering plan the United States and Canada are divided into areas, one or more for each state, each of which is assigned a 3-digit code with a 0 or 1 as the middle digit. Each central office within an area is assigned a code consisting of 2 letters plus 1 digit. The two letters are derived from the initial two letters of the office name and the digit is assigned arbitrarily from 1 to 9. The combination of the 3-digit area code, the 2-letter 1-digit office code, and the 4-digit customer telephone number constitutes the destination code which is to be dialed by operators and customers to reach the called telephone. For calls to points within the home area, the area code is not required. The telephone dial is lettered ABC to WXY, corresponding to digits 2 to 9. By using these eight pulls for the first two letters and the digits 1 to 9 for the third pull, 576 office codes are available theoretically for each area. However, to avoid names which create difficulties in spelling and pronunciation, a standard list containing about 300 names was established from which all new central office names are to be selected and to which existing names are to be converted as rapidly as practicable. Accordingly, 350 to 400 central office codes are the practical maximum for an area, taking into consideration the limited multiple use of the same names, the provision for future growth, and the desirability of natural geographical boundaries.

APPLICATION OF NATION-WIDE NUMBERING PLAN IN PENNSYLVANIA

The application of the nation-wide numbering plan in Pennsylvania is an engineering problem of large magnitude. As the first step studies were made to determine the economical number of numbering plan areas and their boundaries. The larger the number of areas, the easier the assignment problem, but the more onerous the problems which arise from the need for simple instructions for cross-boundary dialing. Table I shows the arrangement finally adopted for the State of Pennsylvania.

Table I. Numbering Plan Areas—Pennsylvania and Delaware

Area	Area Code	Ultimate No. of Codes*
Philadelphia, Eastern Pennsylvania, and Delaware	215	330
Central Pennsylvania	717	381
Northwest Pennsylvania	814	271
Pittsburgh and Western Pennsylvania	412	384
Total		1,366

*Includes independent company central offices.

In addition to being included in the 215 area to permit cross-boundary dialing on a 2L+5N basis from Philadelphia and Eastern Pennsylvania, the State of Delaware has been assigned a code of its own, code 302, to be used in the completion of dialed traffic from other parts of the United States.

PROGRESS IN CONVERSION TO ULTIMATE NUMBERING PLAN

The nation-wide numbering plan was definitely formulated in 1949 and the application in Pennsylvania began soon thereafter. Fortunately, the ensuing period was one of rapid growth, which demanded large additions to many central offices and provided opportunities for economical introduction of 2–5 numbering.

In 1949 Philadelphia was the sole 2–5 area. Pittsburgh, Allentown, Bethlehem, Scranton, Wilkes-Barre, Norristown, New Castle, and New Kensington are among the larger exchanges which have been converted since that time. Plans have been initiated for the conversion of the great majority of the remaining offices to obtain 2–5 numbering for most of the 3,000,000 telephones in Pennsylvania by the end of 1957. Table II shows the progress to date.

Table II. Progress in Conversion to 2-5 Plan in Pennsylvania

Jan. 1	Telephones on 2-5 Basis
1949	704,000
1950	1,059,000
1951	1,210,000
1952	1,417,000
1953	1,630,000
1954	1,847,000

Table III. Progress in Outward Operator Dialing in Pennsylvania

No. of Toll Operating Centers		No. Converted for Dialing	
1954	Ultimate	Equipped with dials	Equipped with Key Pulsing
111	63	66	23

STANDARD CENTRAL OFFICE NAMES

While the major effort in seeking adherence to the nation-wide numbering plan has been in the direction of changing to 2–5 numbering, the matter of the central office names has been given a great deal of attention. Beginning in 1951, offices converted from manual to automatic operation and automatic offices converted to 2–5 numbering have been assigned central office names from the standard list. From the standpoint of the telephone company the conversions involve only technical changes: recasting the slugs for printing the directory, changing the telephone number card, changes in the records, etc. For the customers it involves a profound change in their telephone calling habits, and a great deal of instruction and persuasion is necessary to obtain a satisfactory acceptance. A fundamental plan has been made for the eventual conversion of all the existing central office names in Pennsylvania. Up to the present 129 central office units out of a total of 590 have been named from the standard list.

Equipment at Terminals

EQUIPMENT AT CALLING END—OPERATOR DIALING

With operator dialing, the equipment at the calling end includes a means for sending out the signals corresponding to the digits required for the completion of the call. While the term "dialing" has been used to describe this process, actually it is quite common to substitute for the dial a keyboard, similar to that used in an adding machine, which is capable of sending ten combinations of two tones, one for each digit. There are many advantages to be obtained from the multifrequency pulsing scheme, including a saving of operating time and toll-line time. For new switchboards and for converting existing manual switchboards, multifrequency key pulsing equipment is more economical than dial equipment; for switchboards equipped at present with dials, conversion to key pulsing may be economical if a large proportion of the originating traffic is to offices equipped to receive multifrequency pulses.

Table III shows the progress of outward operator dialing in Pennsylvania. The number of toll centers in the ultimate as well as the present has been shown to give a true picture, since discontinued operating centers will not usually be converted.

EQUIPMENT AT CALLING END—CUSTOMER DIALING

The dial of the telephone instrument is satisfactory for sending out the pulses necessary for customer dialing of toll calls. The main problem at the calling end is to supply a means for automatic timing and charging. The recently developed crossbar central offices are equipped with locally installed common control apparatus which accomplish this purpose, using coded perforated tapes which are processed at an automatic accounting center to bill the customers. For the step-by-step and panel central offices, centralized tape-perforating equipment has been developed to time and charge for customer-dialed toll calls. For the present, identification of the calling customer is accomplished by calling in an operator who obtains the calling number verbally and pulses it into the charging equipment.

At present there are 19 central offices in Pennsylvania equipped for local automatic charging of customer-dialed calls, but the telephone users in all but one of these offices are dialing toll calls to a restricted area. These offices can be converted to nation-wide customer dialing at a comparatively small investment per office, but it is not practicable at this time to tie them into the nation-wide network. Thus far only the Valley central office, comprising 13,000 telephones, has been equipped, and the customers have been dialing to points throughout the nation since December 1953 with satisfactory results.

Plans have been completed to introduce restricted customer dialing on the centralized charging basis in Philadelphia beginning in 1954 and in Pittsburgh in 1955. This method of operation will be extended to the other central offices as rapidly as warranted by engineering and service considerations. The geographical area covered will be extended subsequently to embrace the entire country.

EQUIPMENT AT INCOMING END

The equipment which has been provided at the incoming offices to complete local dialed calls may also be used for completing operator- and customer-dialed toll calls, although a separate toll-completing train is usually provided in the larger offices. On the other hand, the incoming toll selectors in step-by-step offices and their counterparts in panel and crossbar equipment, which must be installed at the toll centers to receive operator- and customer-dialing pulses, constitute one of the principal items of cost in the toll-dialing system. The progress in Pennsylvania to date on inward toll dialing is shown in Table IV.

Toll-Switching Points

GENERAL REQUIREMENTS

A large percentage of the toll calls involve connections on which it is uneconomical to provide toll lines which directly connect the originating and terminating toll centers. A system of toll-switching centers must be provided which is in economic balance with the toll-line network, resulting in a minimum combined cost for outside plant and equipment. A general Bell System toll-switching plan had been established which was satisfactory for manual operation. For dial operation two steps were taken on a Bell System basis, development of the switching equipment, and establishment of a toll-switching plan suitable for both operator and customer dialing.

The application in Pennsylvania is of special interest from several standpoints: Philadelphia was the site of the first crossbar toll-switching system, installed in 1943; the first system incorporating the newly developed card translator, which is an essential requirement for routing toll calls dialed on a destination code basis, was installed in Scranton in 1953; with the completion of the Harrisburg equipment in 1955 Pennsylvania will be the first state, requiring more than one switching point, to complete its program.

DIAL TOLL-SWITCHING PLAN IN PENNSYLVANIA

The dial toll-switching plan in Pennsylvania is shown graphically on the map of Fig. 1. The principal switching points are Philadelphia, Pittsburgh, Scranton, and Harrisburg. The map also shows the two subsidiary switching points at Altoona and Warren which it is planned

Table IV. Progress in Inward Toll Dialing in Pennsylvania

Number of Inward Toll Centers			Number of Toll Lines, 1954	
1954	Ultimate	Equipped for Inward Dialing	Total	Equipped for Dialing
78	48	35	16,000	11,400

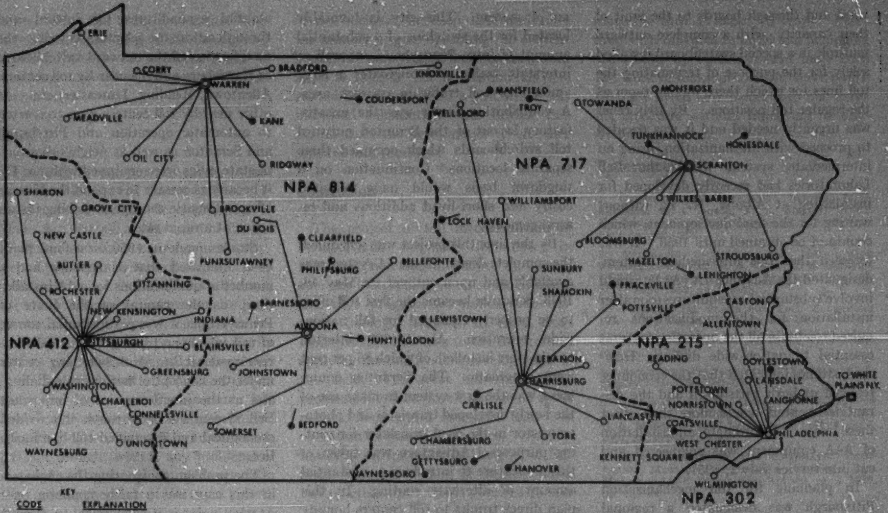

ERIE
CORRY
WARREN
BRADFORD
KNOXVILLE
MEADVILLE
KANE
COUDERSPORT
MANSFIELD
TROY
WELLSBORO
TOWANDA
MONTROSE
TUNKHANNOCK
HONESDALE
OIL CITY
RIDGWAY
NPA 814
NPA 717
SCRANTON
SHARON
GROVE CITY
BROOKVILLE
DU BOIS
WILLIAMSPORT
WILKES BARRE
NEW CASTLE
CLEARFIELD
LOCK HAVEN
BLOOMSBURG
BUTLER
PUNXSUTAWNEY
PHILIPSBURG
BELLEFONTE
SUNBURY
HAZELTON
STROUDSBURG
LENIGHTON
ROCHESTER
KITTANNING
SHAMOKIN
FRACKVILLE
POTTSVILLE
EASTON
NEW KENSINGTON
INDIANA
BARNESBORO
LEWISTOWN
ALLENTOWN
NPA 412
PITTSBURGH
BLAIRSVILLE
ALTOONA
HUNTINGDON
NPA 215
READING
DOYLESTOWN
LEBANON
LANSDALE
TO WHITE PLAINS N.Y.
GREENSBURG
JOHNSTOWN
HARRISBURG
POTTSTOWN
WASHINGTON
NORRISTOWN
LANGHORNE
CHARLEROI
CARLISLE
CONNELLSVILLE
LANCASTER
COATSVILLE
SOMERSET
BEDFORD
CHAMBERSBURG
YORK
PHILADELPHIA
WEST CHESTER
UNIONTOWN
WAYNESBURG
GETTYSBURG
HANOVER
KENNETT SQUARE
WAYNESBORO
WILMINGTON
NPA 302
KEY
CODE
EXPLANATION
REGIONAL CENTER
SECTIONAL CENTER
PRIMARY OUTLET
TOLL CENTER
DISCONTINUED TOLL CENTER
NUMBERING PLAN AREA BOUNDARY
CIRCUIT GROUP TO HOME CENTER

Fig. 1. Toll switching points, homing arrangements, and numbering plan boundaries for state of Pennsylvania

to establish around 1957. The ultimate toll centers, which are also centralized automatic message accounting centers, are shown and their "home" switching points are indicated. Numbering plan areas, which have been made to coincide with the boundaries of the homing areas, have been added to complete the picture. The salient features of each of the switching points will be described in chronological order.

TOLL-SWITCHING EQUIPMENT AT PHILADELPHIA

In the early 1940's the problem of toll switching in Philadelphia, aggravated by the large increase in toll traffic caused by the war effort, became very acute. After considerable study it was decided to install the newly developed no. 4 toll-switching equipment which employed as its basic element the crossbar switch, already used successfully in local equipment. Another important feature borrowed from the local equipment was the "marker" which concentrated within it the functions of control required for connection. The installation was completed in the summer of 1943, the cutover taking place on August 21 of that year.

This first edition of the no. 4 system had one piece of equipment which was never needed again. Because it was installed at a time when toll-line dialing was comparatively new, it was necessary to incorporate into the system a switchboard which would do the dialing for the distant ringdown points. The near-by outward switchboards had been converted to dialing prior to the no. 4 cutover, but the more distant points did not have sufficient traffic to and through Philadelphia to warrant the large expenditures involved. Subsequent installations beginning with the New York no. 4 in 1948 did not require this switchboard to obtain a substantial load for the automatic equipment and the traffic from ringdown points continued to be handled at the old manual inward and through switchboards.

The initial no. 4 project provided terminations for 2,200 intertoll trunks, of which 55 per cent were dial and the remainder ringdown. In 1943 dial trunks were limited in length to 250 miles because of the inability of the toll-line facilities to transmit dial pulses or supervisory signals beyond this range. While multifrequency receiving equipment, installed here for the first time, removed the pulsing limitation, the supervision limitation still remained until the development of voice-frequency signaling equipment 4 years later. Under these conditions Richmond, Va., was the furthest dial point.

The Philadelphia no. 4 system will reach its maximum capacity in 1954 with the installation of the seventh addition. The system has fufilled all the expectations of its designers with respect to both service and economy. For 5 years, because of governmental limitations on Bell System production, it was the sole example of its kind. Today it has 25 companions in the Bell System, with many more to come. A comparison of its present and 1943 status is shown in Table V.

Toll-Switching Equipment at Pittsburgh

In 1949 a crossbar tandem system was installed at Pittsburgh for switching short-haul toll traffic on a dial basis. With this exception, inward, outward, and through toll traffic was handled with the manual ringdown method at the downtown toll switchboards. Intertoll trunks were terminated in the downtown out-

Table V. Status of Philadelphia No. 4 System

	At Cutover Aug. 21, 1943	1954
Total number of trunks	2,229	4,050
Number of dial trunks	868	1,660
Number of multifrequency trunks	194	1,750
Number of messages handled per day	60,000	156,000
Number of switchboard positions required	32	11

ward and through boards to the limit of their capacity, with a complete outward multiple in a special switchboard installed solely for the purpose of terminating the toll lines for which there was no room at the regular toll positions. By 1949 relief was urgently needed and it was decided to proceed with mechanization, using an intermediate system which the Bell Laboratories had recently developed for installation at Albany, N. Y., without waiting for the final development, which could not be obtained until 1953. It was foreseen that this intermediate system, designated the Advance 4A (A4A), would involve substantial expenditures soon after installation for the modifications required to obtain all the operating features essential to nation-wide dialing. However, studies indicated that the economies realized at both Pittsburgh and the distant toll centers would more than offset these losses. Accordingly, an installation of A4A equipment was engineered and cut into service July 1, 1951.

In planning for toll mechanization Pittsburgh was designated a regional center to serve as the major switching point for a large portion of the traffic between the East and the West. The development of voice-frequency signaling since the Philadelphia project permitted the establishment of dial intertoll trunks to all major switching points in the country. The A4A was equipped for about 3,000 intertoll trunks of which two-thirds were dial-operated and appeared only on the crossbar system. The other one-third, which remained ringdown, was left on the downtown switchboards and multipled to the outgoing crossbar switches for outward operation. The inward and through toll boards were combined and reduced to serve the remaining incoming ringdown traffic. The outward toll switchboards were modified for multifrequency key-pulsing.

The Pittsburgh system has grown rapidly since its cutover in 1951. Toll growth has been high; as an additional factor, a substantial amount of through switching was transferred to it from other manual points now that the capacity was available. It became apparent that conversion to 4A, with its full nation-wide switching features, should not be postponed. Accordingly, this was combined with the engineering of the second addition and completed in March 1954. The office now has grown from 3,000 to 4,200 intertoll trunks.

Toll-Switching Equipment at Scranton

Scranton was selected as the third location in Pennsylvania for installation of a no. 4 system. This city is favorably located for the switching of a substantial amount of intra-Pennsylvania as well as interstate traffic and originates a large amount of toll traffic in its own area. A contributing factor was the unsatisfactory layout of the Scranton outward toll switchboards which occupied three separate locations. Continuation on a ringdown basis would have required costly and short-lived additions and rearrangements.

By the time this project was scheduled, the complete design for the 4A system was available and upon cutover on May 16, 1953, Scranton became the first toll office to be properly equipped for full nation-wide operation. About 900 intertoll trunks were installed, of which 85 per cent were automatic. The Scranton equipment was the first system to make use of the newly developed transistor and phototransistor in the card translator for routing purposes. Advantage was taken of this 4A feature to introduce a substantial amount of alternate routing. In this plan direct trunks to toll centers homing elsewhere are established on a high-usage basis, with the overflow being routed over liberally engineered trunk groups via the home switching point. Similarly, the toll centers homing on Scranton may have high-usage groups from distant 4A centers with the overflow routed via the Scranton equipment. By 1953, transmission techniques had improved to the point where delayed through switching for the 4-wire circuits could be handled satisfactorily with 2-wire cords, and toll-line provision had become so liberal that the amount of delayed switched traffic had been substantially reduced. Accordingly, the dial conversion was accompanied by the replacement of the existing switchboards with a new, modern board equipped with key pulsing and designed to handle the delayed through traffic as well as the remaining ringdown traffic. This is a departure from the procedure followed at both Philadelphia and Pittsburgh, where the old switchboards were left in place and modified for outward dialing and where especially designed new switchboards, using 4-wire cords, were installed to handle delayed through traffic.

Toll-Switching Equipment at Harrisburg

Even prior to the installation of the Philadelphia no. 4 system in 1943, a step-by-step toll train had been established in Harrisburg for switching between toll centers in central Pennsylvania. This was the first venture into intertoll dialing in Pennsylvania and involved substantial expenditures for inward and through automatic selectors as well as the modification of the outward switchboards at Harrisburg and the near-by toll centers: Allentown, Reading, Lancaster, etc. As other near-by toll centers were converted to automatic operation and Pittsburgh and Scranton as well as neighboring out-of-state cities were equipped with no. 4's, it became necessary to expand the Harrisburg automatic through-switching installation at a rapid rate.

Studies made in 1950 on a fluid plant basis indicated that Harrisburg had a number of qualifications for a 4A installation: as the capital of the State of Pennsylvania it was a substantial source of toll telephone traffic; its central location made it an ideal switching center under the concept of nation-wide dialing; and as the junction point of important Bell System toll cable routes it provided economical and diversified toll-line facilities.

The problem confronting the engineers in this case was a fairly common one: At what point is it economical to discard working equipment giving satisfactory service in favor of a new instrumentality whose coming appears inevitable? It was complicated by the need to determine the effect of the conversion at Harrisburg on the 90 offices which connect with it. The economies in operation due to supplanting the ringdown method were not available in this case. Nevertheless, the advantages of common control equipment, alternate routing, multifrequency pulsing, and other features of the 4A equipment, coupled with the large cost involved in expanding the step-by-step installation to meet the expected growth, resulted in a decision to proceed with the installation of a 4A system at Harrisburg for service on December 1, 1955.

In engineering the Harrisburg 4A system every effort was made to conserve the investment in the existing equipment. Inward toll lines which carry only terminal traffic to the Harrisburg area will remain on the step-by-step selectors. Similarly, some outward terminal trunks will be retained in the switchboard multiple. The existing toll switchboard at Harrisburg wll be left in place, except that the present dial-pulse equipment will be converted to key pulsing.

Floor space of the proper strength and height is available in the present building on floors now occupied by office forces. The initial design will care for 1,500 toll lines of which 88 per cent will be automatic and the remainder ringdown. The layout is engineered to permit a threefold increase without expensive rearrangements. This has been done to facilitate the eventual upgrading of Harrisburg from its initial status to a higher rank of switching center, if future conditions warrant it.

Costs

The Bell System investment in equipment installed for toll dialing in Pennsylvania to date is on the order of $35,000,000. About $12,000,000 are invested in the no. 4 systems and the remainder in step-by-step equipment, new switchboards and modifications of existing installations, automatic message accounting equipments, signaling circuits, etc. Each installation was justified on its own merits and as part of the over-all arrangement from the standpoint of service improvement and economy. The economy has come mainly from the reduction in traffic-operating labor, yet toll mechanization has not brought with it any problems of disposal of excess operating forces. On the contrary, because of the rapid growth of the business, between 1941 and 1954 the number of girls operating toll switchboards in Pennsylvania has increased from 5,000 to 10,500. Mechanization has permitted better service at a cost which has increased substantially less during this period than the great majority of other services and products.

Conclusion

The project of toll mechanization in Pennsylvania, originally formulated in 1941, has increased in quality and scope over the intervening years to its culmination as a program for nation-wide customer dialing. Throughout these years Pennsylvania has been in the forefront in applying the Bell System developments in this field. The great majority of the problems involved have been solved or are near solution and a successful trial is in operation at one central office. No date has been set for reaching the goal of full-scale customer dialing in Pennsylvania, but no great optimism is required to estimate the time at about 10 years. Planning with this goal in mind and achieving the results up to the present have been the work of many people who have painstakingly fitted new pieces into the existing structure without impairing its day-by-day operation.

References

1. Crossbar Toll Switching System, L. G. Abraham, A. J. Busch, F. F. Shipley. *AIEE Transactions (Electrical Engineering)*, vol. 63, June 1944, pp. 302–09.

2. Automatic Switching for Nation-wide Telephone Service, A. B. Clark, H. S. Osborne. *AIEE Transactions (Electrical Engineering)*, vol. 71, pt. I; Sept. 1952, pp. 245–48.

3. Fundamental Plans for Toll Telephone Plant, J. J. Pilliod. *Ibid.*, pp. 248–56.

4. Nation-wide Numbering Plan, W. H. Nunn. *Ibid.*, pp. 257–60.

5. Automatic Toll Switching Systems, F. F. Shipley. *Ibid.*, pp. 261–68.

6. Application of Multifrequency Pulsing in Switching, C. A. Dahlbom, A. W. Horton, Jr., D. L. Moody. *AIEE Transactions*, vol. 68, pt. I, 1949, pp. 392–96.

7. Fundamentals of the Automatic Telephone Message Accounting System, John Meszar. *AIEE Transactions*, vol. 69, pt. I, 1950, pp. 255–69.

No Discussion

Precision High-Current Computer Power Supplies

ALLEN B. ROSENSTEIN
ASSOCIATE MEMBER AIEE

THE increasing size and complexity of analogue and digital computers has resulted in an ascending demand for expanding quantities of well-regulated and filtered d-c power. As the machines have grown in size the regulated power demands have gone from milliamperes to amperes, to tens of amperes, and now to hundreds of amperes. These increased loads have been supplied by various methods, most of which offered compromises with desired performance.

With the present-day trend toward computer packaging, component miniaturization, and the intensified search for reliable maintenance-free components, the power supply has often proved to be the weakest link in the system. D-c sources which are adequate at a few amperes are swamped at tens of amperes, while conventional high-current supplies have excessive bulk or maintenance.

This paper describes high-current precision d-c power supplies which have been developed to meet the requirements of modern computers. The power supplies are unique in that for the first time a high-current magnetic-amplifier-controlled selenium rectifier has been produced with the response and electrical performance of the conventional electronic supply and yet possessing the mechanical advantages and life of the magnetic-amplifier–selenium rectifier. Results are presented of tests conducted upon supplies constructed for the large model-*107* digital computer developed by the Computer Research Corporation for the U. S. Bureau of Aeronautics. This computer is to serve as a central clearing house and perpetual inventory for all of the spare parts of the Bureau.

Power Supply Requirements

The basic computer power supply requirements are economic. The cost of "down time" is frequently prohibitive. The need for complete dependability and freedom from any maintenance or service is ever present and ever increasing. The ideal power supply would be a sealed unit, completely static, with a service-free life of 10 to 20 years. The power supplies to be described approach these criteria and are expected to require little more than yearly servicing.

With increasing frequency, costs are dictating mechanical requirements. The trend in mechanical computer design is toward compact, high-performance machines capable of installation in an office or moderate-sized laboratory. Here savings in power supply size and weight can be translated directly into greater equipment utility and reduced assembly costs. Many civilian and military computers are being used in mobile or air-borne applications. Here, of course, size and weight are of primary importance.

Further mechanical requirements in-

Paper 54-313, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted October 27, 1952; made available for printing April 26, 1954.

Allen B. Rosenstein is with the University of California, Los Angeles, Calif., and a consultant to INET, Los Angeles, Calif.

Table I. Physical and Electrical Characteristics of 225-Volt 15-Ampere Precision Power Supplies

	60-Cycle, Self-Cooled	400-Cycle, Forced-Cooled
Steady-state regulation, per cent	more than ±0.1	more than ±0.1
Dynamic regulation, per cent	more than ±0.15	more than ±0.15
Ripple, per cent rms	less than 0.1	less than 0.1
Efficiency, per cent	75	71
Volume, cubic feet	5.9	1.5
Weight, pounds	369	92

clude insensitivity to temperature, shock vibration, etc. Mobile equipment especially, and office machines to some extent, must be able to withstand some degree of vibration and shock. Military installations will often experience continuous high vibration and shock.

Although temperature-controlled space can often be made available for the computer power supply, it is an added expense and inconvenience. A power supply which can operate over a nominal ambient temperature range has increased utility. A supply capable of satisfactory performance with extremes of temperature, altitude, etc., is suitable for military or civilian applications at desert, tropical, or arctic posts. The ability of equipment to continue in operation at remote locations with virtually no maintenance or replacement of parts will make possible increasingly wider application of computers.

The electrical requirements of a computer power supply are dictated by the characteristics of the machine load. Although these requirements will vary from machine to machine, a reasonably uniform electrical specification, can be drawn, depending upon the type of machine and its function. Since over-all machine performance is of primary importance, the high-voltage d-c supplies are often designed to give a slow initial voltage build-up to minimize the starting shock upon the computer components.

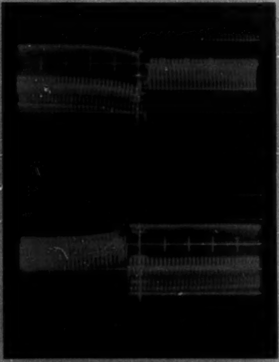


Fig. 1. Dynamic load regulation; 40-per-cent step change in load. Each major division = 0.25 volt

Top—120 per cent/80 per cent
Bottom—80 per cent/120 per cent

Full input to output efficiency of larger power supplies is receiving close scrutiny by computer designers. Many computers are being installed in offices where the power service is limited. Practically all of the power supplied by the computer's d-c source is ultimately converted into heat within the computer itself. It can be seen that the "furnace value" of a computer will vary directly as the output of the power supply and inversely as the power supply efficiency. The energy dissipated in modern high-speed computers will often overload the existing ventilating capacity of an area. To allow operation in restricted office space and also to provide optimum working conditions for the computer's components, computers are being built with integral air-conditioning systems. The corelation between power supply efficiency and air-conditioning cost is obvious. Designers are now calling for power supplies with over-all efficiencies of 70 to 75 per cent or better.

Since digital computers are pulse-operated, the restrictions upon all forms of deviation of the power supply's d-c output potential are severe. A pulse, regardless of its origin, can possibly cause a false count within the computer.

To design the proper supply, its electrical output specifications must carefully specify the required voltage response under all types of loading and with maximum variations in the a-c power input. Not only must the steady-state regulation be given but ripple must be considered as an addition to the instantaneous dynamic regulation. Often the maximum allowable instantaneous voltage deviation will be specified for the worst possible combination of a-c line variation and ripple, with loading consisting of pulses from 0 cycle per second to several megacycles. Steady state regulation, dynamic regulation, and ripple are each commonly specified to be held within from 0.1 to 1.0 per cent.

Steady-state regulation is a direct function of the closed-loop system gain. Ripple is controlled by filtering. The instantaneous dynamic regulation will be determined by the impedance of the supply. Because of the broad band of loading frequencies to be accommodated, it presents, generally speaking, the more difficult design problem.

Description of Computer Power Supply Systems

Each of the following systems requires a flywheel or shunt energy storage element to ensure adequate dynamic regulation. The natural response and internal impedance of the power source determine the magnitude of this stored energy.

Battery banks were an early source of high-current d-c computer power. A battery's energy storage capacity is high, and its dynamic resistance is low. Filtering requirements are, of course, small. However, maintenance, bulk, and first cost are relatively high. In addition, a well-regulated charging source is required to hold the battery potential at its required level.

For low current requirements of 1 or 2 amperes or less, the fully electronic supply with hard vacuum series-regulating tubes has been extensively applied. Regulation can be made excellent, response is extremely fast, and internal impedance is

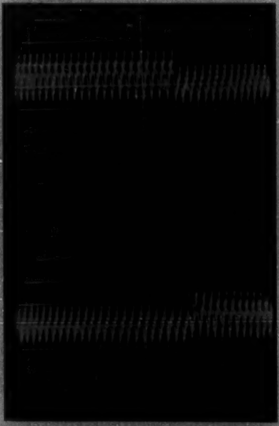


Fig. 2. Dynamic line regulation at 100-per-cent load; 10-per-cent step change in a-c supply voltage. Each major division = 0.25 volt

Top—240 volts/220 volts
Bottom—220 volts/240 volts

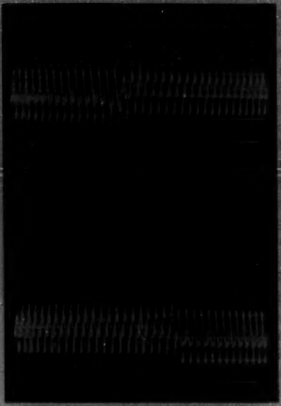


Fig. 3. Dynamic load regulation; 50-per-cent step change in load. Each major division = 0.25 volt

Top—50 per cent/100 per cent
Bottom—100 per cent/50 per cent

relatively low. The shunt energy storage requirement is correspondingly small. The series regulator has the disadvantages of low efficiency, fairly high bulk, and high maintenance potential. Since the full output current is carried by vacuum tubes, this system has a large number of tubes with the corresponding possibility of failure.

Precise high-current d-c computer sources have generally utilized grid-controlled thyratrons. The efficiency of the thyratron power supply is fairly good for the higher voltage. The internal impedance of the power service is relatively low. Maximum response is between 1/2 and 1/6 cycle, dependent upon whether the system is single-phase full wave or 3-phase full wave. The shunt stored energy is greater than that required for a comparable series regulator. Efficiency at lower voltages (below 200 volts) drops rapidly. The equipment is relatively bulky. With two to six thyratrons, the supply can be expected to require tube service. The gas-filled tubes are not readily portable and are sensitive to shock, vibration, and temperature

While the electrical performance of the thyratron power supply can be made quite satisfactory for high-current computer applications, its mechanical characteristics leave much to be desired. The desired high-current electrical and mechanical characteristics have been achieved in the power supply to be described in this article by substituting the magnetic-amplifier-selenium rectifier for the thyratrons

High-Speed Magnetic-Amplifier Performance

The magnetic-amplifier-regulated selenium rectifier achieves the full mechanical criteria for computer power supplies. The power section is completely static and composed of highly dependable elements of extremely long life. The components are capable of withstanding even a high shock test and prolonged vibration. The over-all bulk is at practically the irreducible minimum consistent with today's materials. The major weight component is the power transformer. Temperature characteristics are excellent and temperature extremes ranging from below 0 to +160 degrees Fahrenheit can be met.

The major drawback to the application of magnetic amplifiers, i.e., nonlinear reactors, has been the popular concept that magnetic amplifiers are extremely slow responding devices. This concept is correct only in a very limited relative sense. Since magnetic-amplifier performance is dependent upon its carrier frequency, it is correct to say the magnetic-amplifier response is slower than that of vacuum-tube amplifiers. However, our knowledge of the mechanics of magnetic-amplifier operation has advanced to the point that magnetic-amplifier response is limited only by its power supply frequency. A circuit featuring 100-per-cent half-cycle response has already been demonstrated by Ramey.[1,2] The equipment described in this article utilizes a recently developed circuit capable of producing 100-per-cent response of 1/6 cycle or better. This is performance identical to that of the best thyratron system where again response is limited by carrier frequency.

Table I and Figs. 1 through 4 give the physical and electrical characteristics of a small computer power supply rated 225 volts, 15 amperes. The steady-state regulation of the supply is ±0.1 per cent while the dynamic load regulation with ±20-per-cent load changes does not exceed ±0.15 per cent. It is to be noted from the oscilllographs, Figs. 1 through 4, that the dynamic regulation due to a step change in load or a-c supply is practically the same as the steady-state regulation. As the response of the power section is only 1/6 cycle this small difference between dynamic and steady-state regulation for nominal load changes has been found to be in the damping circuits of the preamplifier rather than in the power section.

To obtain a ripple of less than 0.1 per cent rms conveniently, a relatively large shunt energy storage component has been used. As a result of this, excellent dynamic regulation under even extreme changes in loading are obtained. Fig. 3 shows the output voltage deviation with instantaneous load changes between 50 and 100 per cent. The dynamic regulation is still less than 0.10 per cent.

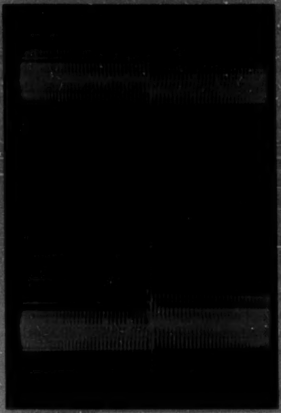


Fig. 4. Dynamic line regulation at 100-per-cent load. 10-per-cent step change in A-C supply voltage. Each major division = 0.25 volt

Top—240 volts/220 volts
Bottom—220 volts/240 volts

It is interesting to observe that in this power supply a balanced design has been achieved. The dynamic response under normal operating conditions does not exceed the peak value of the ripple.

The efficiencies and 60-cycle mechanical specifications listed in Table I are for a normal self-cooled design. The 75-per-cent efficiency presents a compromise with the available space and cost. By increasing the power section component size a nominal amount, it is possible to bring the over-all efficiency of the power supply up to 80 to 85 per cent.

Where space and weight become the major criteria, some advantage can be gained with forced air cooling. However, major mechanical and electrical advantages are gained by utilizing higher supply frequencies of, say, 400 cycles or more. Response time varies inversely with frequency with the response of a 60-cycle supply equal to $1/(6\times 60)=0.0028$ second while the response at 400 cycles becomes 0.00042 second. As the weight of a precision power supply is concentrated in its iron-cored elements, its weight will also vary almost inversely with frequency.

Table I compares the physical characteristics of a self-cooled 60-cycle design with a forced-air-cooled 400-cycle design where the electrical performance is left unchanged. The reduction in weight and volume is of major interest to the computer designer. While some of the reduction in volume came from the reduced size of the iron-core elements, a substantial volume saving came from the reduction in shunt energy storage required for the 400-cycle supply. With the higher frequency, filter requirements are naturally reduced, as the improved response takes care of the subharmonic load changes. In view of the performance gains and weight and volume savings effected with the higher supply frequencies, it is expected that more and more computers will be built using 400-cycle or higher a-c power.

Design Considerations

The fundamentals of regulated precision power supply design revolve around the restrictions laid down by the basic "three R's" of ripple, regulation, and response. Although the designer might at first consider these to be three independent variables, he soon finds them to be mutually dependent. Full establishment of two of the variables usually results in a determination of the third. The block diagram of the precision computer power supply is quite simple; see Fig. 5. Alternating voltage is supplied to the magnetic-amplifier–selenium rectifier where it is reduced to the desired level and rectified. The raw direct current from the rectifier is smoothed by the filter, which contains a shunt energy storage component. The filtered d-c output is compared with a suitable standard and the difference amplified by the preamplifier to control the magnetic amplifier.

Although it is common practice to speak of steady-state regulation and dynamic regulation, it is recognized that this is only distinguishing between the zero frequency error and the frequency dependent error caused by a variable frequency load. Within the long-term stability limits of the reference and the requirements for closed-loop stability, almost any desired steady-state regulation can be achieved by increasing the closed-loop gain. The closed-loop diagram of the power supply is given in Fig. 6. The diagram indicates that errors or deviations in output voltage result from changing loads and variations in the a-c supply.

The frequency response (dynamic regulation) of the system for changes in both a-c supply and d-c load can be predicted from the equivalent circuit of Fig. 7. Since most of the changes in the d-c loading upon the power supply will be caused by variations in tube plate currents of the computer load, it is natural to represent the load by the conventional equivalent circuit of a tube, i.e., a resistance R_g in series with a constant a-c generator $e_1(t)$.

The magnetic-amplifier–selenium rectifier has been represented by an adjustable d-c source E_s and its internal resistance R_I. At this point the somewhat discontinuous nature of the voltage output of the magnetic amplifier should be recognized. The magnetic amplifier is not being used as variable reactance but as a magnetic gate with an alternating voltage gating action entirely similar to that of the phase-controlled thyratron. As in thyratron action, the first portion of the alternating wave is absorbed by the magnetic amplifier before its reactor fires (saturates) and allows the rest of the alternating wave to be applied to the rectifiers. The d-c output of the rectifier is controlled by varying the amount of the a-c wave blocked by the magnetic gate. By proper choice of materials the reactance of the reactors after saturation can be made negligible at the carrier frequency. The magnetic-amplifier–selenium rectifier then can be represented by a gated d-c source (with output wave form as shown) and its internal resistance.

The magnetic amplifier has a response of 1/6 cycle for the polyphase connection. With a 60-cycle carrier, signals above 360 cycles per second are attenuated. For frequencies in the even low kilocycles, the magnetic amplifier is, of course, completely insensitive. To maintain the output potential deviation within required limits under high-frequency loading, it becomes necessary to supply a shunt energy storage element or low-pass filter. A variety of circuits and elements can be used. For simplicity of analysis, a capacitor C has been used in the equivalent circuit.

At the higher frequencies, R_I will be large compared to X_{cs}. The e_L will be attenuated by the resistor R_L. The component of the high-frequency voltage appearing across the output terminals will be given by

$$\Delta e_0(\omega)=\frac{-jX_c(\omega)}{R_L-jX_c(\omega)}e_L(\omega)$$

Since X_c is frequency variant, the deviation of e_0 is dependent upon both frequency e_L and the magnitude of R_L which is in turn determined by the d-c component of load current.

The low-frequency response study of the system must be modified in light of the voltage gating action of the magnetic amplifier. If the full alternating potential available to the magnetic amplifier were applied, instead, directly to the selenium rectifier, the d-c output voltage

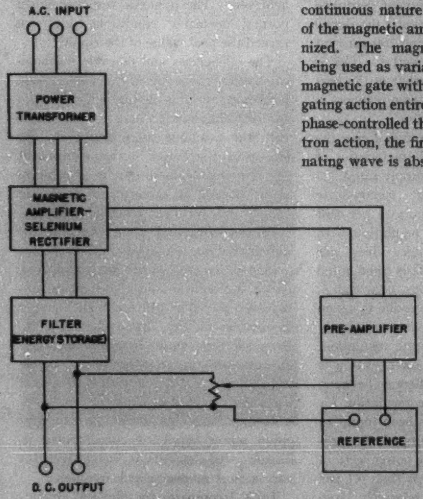


Fig. 5. Sytems block diagram

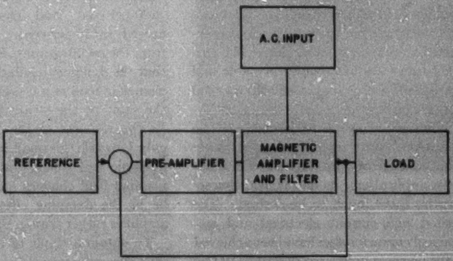


Fig. 6. Closed-loop diagram

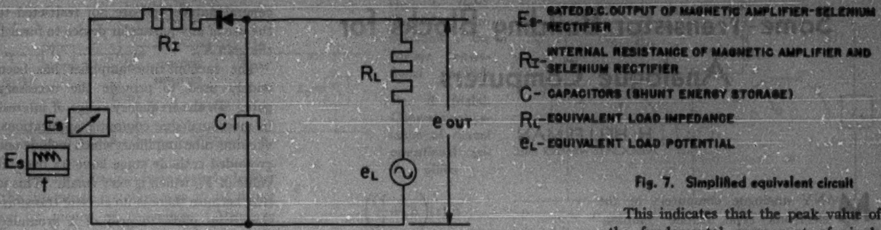

R_I
E_s
C
R_L
e_{OUT}
e_L
E_s- GATED D.C. OUTPUT OF MAGNETIC AMPLIFIER-SELENIUM RECTIFIER
R_I- INTERNAL RESISTANCE OF MAGNETIC AMPLIFIER AND SELENIUM RECTIFIER
C- CAPACITORS (SHUNT ENERGY STORAGE)
R_L-EQUIVALENT LOAD IMPEDANCE
e_L- EQUIVALENT LOAD POTENTIAL

Fig. 7. Simplified equivalent circuit

would rise to a maximum value. It is the function of the magnetic amplifier to block the first part of the a-c wave long enough to reduce the resultant rectified potential to the desired value. The magnetic gate, similar to the electronic gate, no longer has any control over the voltage passed once the gate has fired or is opened. Control can only be effected through variation of the gating time of the succeeding voltage wave which appears 1/6 cycle later. Maximum voltage deviation would then occur if the load were to be switched at the instant that the magnetic amplifier had fired.

To determine this deviation, consider $e_L=0$ and vary the load by switching a resistor R_1 in parallel with R_L. If E_s is assumed to have a constant average value during this interval, then the output voltage C will start to decay exponentially from

$$\frac{R_1}{R_L+R_s}E_s \text{ down to } \frac{R_{11}}{R_{11}+R_s}E_s$$

where $R_{11}=(R_LR_1)/(R_L+R_1)$.

The output voltage deviation E_0 for a period Δt will be given by

$$\Delta E_0=\left(\frac{R_L}{R_L+R_s}-\frac{R_{11}}{R_{11}+R_s}\right)\left(1-\frac{-\Delta t}{\epsilon R_{111}c}\right)E_s$$

where $R_{111}=(R_{11}R_s)/(R_{11}+R_s)$ (i.e., the parallel impedance of R_1, R_L, and R_s).

The role of the internal supply resistance R_s is particularly interesting since it is also a function of the supply efficiency. As R_s becomes smaller, $\Delta E_0\rightarrow 0$ since $R_L/(R_L+R_s)\rightarrow R_{11}/(R_{11}+R_s)$. $R_{111}C$, and therefore the time constant R_{111}, also goes to zero with R_s.

It would seem that every effort should be made to reduce R_s to as close to zero as possible. Unfortunately, upon examining the response of the system to changes in a-c supply voltage, this is found to be undesirable.

For a given gating interval, E_s will vary directly with the a-c line voltage. Should the line voltage drop and cause E_s to drop, the unidirectional nature of the rectifiers will prevent C from discharging through R_s. The discharge time constant remains R_LC, which is a relatively long time. If, however, the line potential were to increase at the instant the gate opens, E_s would start to charge C directly through R_s. Should R_s be reduced to zero, then the variation in E_s would appear directly across the d-c output and it would be impossible to maintain fine dynamic regulation under increases in supply-line potential. A balanced design calling for a given maximum output deviation with rapid a-c line changes must therefore consider the charging time constant of C. The expression for the voltage deviation with a step increase in E_s is

$$\Delta E_0=\frac{R'}{R_s}\left(1-\epsilon^{-\frac{\Delta t}{R'c}}\right)\Delta E_s$$

where $R'=(R_LR_s)/(R_L+R_s)$.

Accurate ripple filter computations can be made by assuming that the system acts as a polyphase rectifier with a capacitor input filter. R_s is further assumed small. Under these assumptions, the current flowing into C will consist of pulses of current which will flow only during a short interval at the peaks of the applied voltage wave. The current flowing into C can then be expressed in terms of a Fourier series

$$i_c(t)=I_{dc}+\sum_{k=1}^{\infty}a_k\cos k\omega t$$

The d-c component is given by

$$I_{dc}=\frac{1}{T}\int_0^T i_c(t)dt$$

The peak fundamental component $a_1=I_1$ is given by

$$I=\frac{2}{T}\int_0^T i_c(t)\cos\omega t dt$$

However, the current only flows into the capacitor during the short peak portion of the cycle, or when the cosine term is approximately unity. During the remainder of the cycle, when the cosine term is less than unity, $i_c(t)=0$; therefore

$$I_1\leq\frac{2}{T}\int_0^T i_c(t)dt=2I_{dc}$$

This indicates that the peak value of the fundamental component of ripple current flowing through C is twice the d-c load current. With the alternating current through C known, the ripple voltage is computed as the product of the alternating current and the capacitor reactance at the fundamental ripple frequency.

Conclusions

Experience has proved the ability of the magnetic amplifier to perform the high-speed, high-current regulating tasks formerly considered feasible only with thyratrons or ignitrons. Heavy current computer power supplies have been built with the electrical performance of the electronic system and the ruggedness and long life of the magnetic amplifier. With reasonable response and ripple specifications there are no practical limitations to the size of the magnetic-amplifier computer power supply.

There appears to be a tendency in American design to use the magnetic amplifier as a new low-level controlling element, while the final power stages of the system are still motor generators, amplidynes, etc. This tendency, if true, overlooks the full potentiality and ultimate application of magnetic amplifiers. INET, Inc., put into service over a year ago a magnetic amplifier with a 160,000-ampere output which was controlled with 500 milliamperes of control current. The response time of the apparatus was less than 1.5 cycles. At the present moment there appear to be no practical restrictions upon the size, voltages, and currents which can be achieved with high-performance magnetic amplifiers.

References

1. On the Control of Magnetic Amplifiers, R. A. Ramey. *AIEE Transactions*, vol, 70, pt. II, 1951, pp. 2124–28.

2. On the Mechanics of Magnetic Amplifier Operation, Robert A. Ramey. *Ibid.*, pp. 1214–23.

3. Vacuum Tube Circuits (book), L. B. Arguimbau. John Wiley and Sons, Inc., New York, N. Y., 1940.

◆

No Discussion

Some Transistor Building Blocks for Analogue Computers

H. HELLERMAN
ASSOCIATE MEMBER AIEE

MANY analogue computers of the electronic type include combinations of a few "building blocks" called operational amplifiers. This paper analyzes the basic properties of voltage and current operational amplifiers and discusses some of the problems of designing the circuitry using junction transistors. Although much has been written about the voltage operational amplifier,[1] especially where vacuum tubes are employed, it is hoped that this presentation of the subject will emphasize those properties of the network configuration which become important if transistors are to be used. A somewhat different approach to making the best use of transistors in the analogue computer type of circuits is to use currents instead of voltages as the dependent variables. This leads to what in some respects seems to be a more reasonable utilization of the natural properties of transistors than the obvious expedient of simply "transistorizing" existing circuits. One of the aims of this paper is to explore this possibility in a rather general way and to relate the principles and philosophy of the current operational amplifier to the more familiar voltage case. The analysis is based on the small signal operation of devices such as vacuum tubes and transistors which allows them to be replaced by the usual equivalent circuits. No attempt will be made to include practical circuit diagrams.

Voltage Operational Amplifier

The Miller voltage operational amplifier can perform the operations of addition, multiplication by a scale factor, integration, and differentiation with an accuracy which can be made to depend upon relatively few circuit components.[2] The basic configuration is shown in rather general form in Fig. 1. Within the large box is a voltage amplifier which, considered as a 4-terminal network, can be specified by the four parameters Y_{11}, Y_{12}, Y_{21}, and Y_{22}. In Appendix I it is shown that the transform of the output voltage designated by $E_o(s)$ is related to the transform of the input voltage $E_i(s)$ by equation 1.[3]

$$E_o(s)=\frac{E_i(s)\left(K_o+\dfrac{Y_f}{Y_o}\right)}{\left(1+\dfrac{Y_i}{Y_s}\right)\left(1+\dfrac{Y_f+Y_L}{Y_o}\right)+\dfrac{Y_fY_L}{Y_sY_o}+\dfrac{Y_f}{Y_s}(1-K_o)+\dfrac{Y_{12}}{Y_s}\left(\dfrac{Y_f}{Y_o}+K_o\right)} \tag{1}$$

where the following notation applies to the amplifier and network within the box

$Y_o=Y_{22}=\left(\dfrac{I_2}{E_o}\right)$: output admittance with 1G short-circuited

$Y_i=Y_{11}=\left(\dfrac{I_1}{E_1}\right)$: input admittance with 0G short-circuited

$K_o=-\left(\dfrac{Y_{21}}{Y_{22}}\right)=\left(\dfrac{E_o}{E_1}\right)$: voltage gain with 0G opened

$Y_{12}=\left(\dfrac{I_1}{E_o}\right)$: ratio of current at 1G to voltage applied at output with 1G short-circuited

If it can be assumed in the frequency range of interest that the following inequalities are substantially satisfied.

$$|K_o|>>\left|\frac{Y_f}{Y_o}\right|$$

$$\left|K_o\frac{Y_f}{Y_s}\right|>>\left|\left(1+\frac{Y_i}{Y_s}\right)\left(1+\frac{Y_f+Y_L}{Y_o}\right)+\frac{Y_f}{Y_s}\left(1+\frac{Y_L}{Y_o}+\frac{Y_{12}}{Y_o}\right)+K_o\frac{Y_{12}}{Y_s}\right|$$

then it can be said that approximately

$$E_o(s)\approx -E_i(s)\,\frac{Y_s}{Y_f}=-E_i(s)\frac{Z_f}{Z_s} \tag{2}$$

It should be noted that these inequalities, which if satisfied make equation 2 valid, are more nearly met if the amplifier K_o has a very high gain and if the input impedance of this amplifier is high and its output impedance is low.

Using equation 2 Table 1 can be constructed, the contents of which are fairly well known. This can be thought of as an application table which indicates the connections of the feedback elements to perform desired operations. Only those combinations in Table I which use resistance and capacitance for Y_f and Y_s have attained widespread use. It should be emphasized that these results are of a general nature and are not restricted to the use of any particular device to furnish the gain K_o.

The vacuum-tube amplifier has been widely used to provide the necessary gain. In the frequency range of interest in most analogue computer applications, vacuum-tube amplifiers which include one grounded cathode stage have an over-all value of Y_{12} which is very small. This is true because there is no electric transmission from plate to grid of a grounded cathode stage, at least at low frequencies. However, this is not the case for transistor amplifier stages. To be perfectly general Y_{12} has been retained in equation 1. In any particular circuit configuration Y_{12} can be determined from its definition. From the second inequality it is necessary that $|Y_{12}|<<|Y_f|$ for equation 2 to hold. The circuit providing the gain may then be described completely in terms of its input and output impedances and its open-circuit voltage gain. It might be worth while to examine some of the characteristics which these transistor voltage amplifiers must have in order to be useful in this application. If simple cascaded transistor stages are to be used, the first general requirement is that they must include an odd number of grounded emitter stages to make K_o negative. It should be noted that Y_{12} for a grounded emitter stage is very small, being of the order of 1 micromho for a typical junction transistor. Although a single grounded emitter stage may be used, it is desirable to add at least one other stage to the grounded emitter to obtain more gain or to keep the input impedance high. For a 2-stage voltage amplifier only a grounded collector or grounded base stage may be added. The grounded collector followed by a grounded emitter combination has the advantage of giving a high input impedance; a value of a few hundred thousand ohms in parallel with the input bias resistor can be achieved. The gain with two high alpha triode junction transistors should be over a hundred. Further cascading to obtain more gain is, of course, possible although the bandwidth deteriorates as more stages are added, and maintaining stability after the feedback is connected becomes a more severe problem.[4]

Paper 54-270, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted October 14, 1953; made available for printing April 30, 1954.

H. HELLERMAN is with Syracuse University, Syracuse, N. Y.

Work on which this paper is based was done under contract with the Rome Air Development Center, U. S. Air Force.

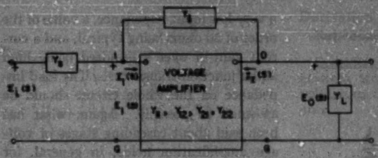
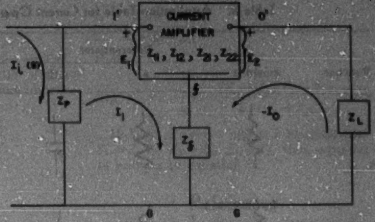


Fig. 1 (left). General configuration of the voltage operational amplifier

Fig. 2 (right). General configuration of current operational amplifier



Thus far it has not been stated whether the interstage coupling was to be direct or by means of capacitors. Although many circuits used as operational amplifiers in analogue computers must be d-c coupled, circuits of this type used in other applications often need not be. The decision as to the kind of coupling rests on the type of input signal on which the circuit must operate. If d-c coupling must be used in a particular application, there are some severe difficulties. One of the hardest obstacles to overcome is drift, especially with temperature changes. Present methods of dealing with the problem involve using special elements which have the property of changing with temperature in such a way that they can be connected in the circuit to compensate for the drift in the amplifier.[8] However, if this is done the amplifier must be constructed with some rather critically matched components. Furthermore, if identical transistors are used in a 2-stage amplifier, only certain combinations will lend themselves to the compensation scheme.

Current Operational Amplifier

As mentioned earlier, another approach to the efficient use of transistors is to use currents instead of voltages as the quantities to be operated upon. To determine the configuration which the current operational amplifier might take, the general principle of duality can be used. If this is done, it is to be expected that a current amplifier will take the place of the voltage amplifier. This current amplifier requires properties which naturally fit the capabilities of transistors. Consider the configuration of Fig. 2. The input is a current $i_1(t)$ and the output a current $i_o(t)$. In terms of the transforms of these currents denoted by $I_1(s)$ and $I_o(s)$ we have equation 3. The derivation of this equation is given in Appendix II, as well as derivations of equations for input and output impedances. (A less general form of equation 3 was suggested independently by J. J. Suran.[9])

$$I_o=\frac{I_1\left(A_c+\dfrac{Z_f}{Z_o}\right)}{\left(1+\dfrac{Z_i}{Z_p}\right)\left(1+\dfrac{Z_f+Z_L}{Z_o}\right)+\dfrac{Z_fZ_L}{Z_pZ_o}+\dfrac{Z_f}{Z_p}(1-A_c)-\dfrac{Z_{12}}{Z_p}\left(\dfrac{Z_f}{Z_o}+A_c\right)} \tag{3}$$

where

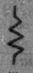
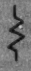
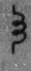
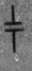
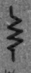
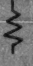
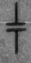
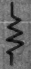
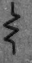
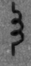
$Z_o=\left(\dfrac{E_2}{-I_o}\right)$: output impedance with $1'f$ opened

$Z_i=\left(\dfrac{E_1}{I_1}\right)$: input impedance with $0'f$ opened

$A_c=\left(\dfrac{I_o}{I_1}\right)$: current gain with output short-circuited

$Z_{12}=\left(\dfrac{E_1}{-I_o}\right)$: ratio of voltage across $1'f$ to the current applied to $0'f$ terminals when $1'f$ is opened

Table I. Application Table for Voltage Operational Amplifier

		Differentiator		Integrator	
Function	Amplifier	a	b	a	b
Y_f..........					
Y_s..........					
$E_o(s)$........	$-\dfrac{R_f}{R_s}E(s)$........	$-RCsE(s)$.......	$-\dfrac{L}{R}sE(s)$......	$-\dfrac{1}{RCs}E(s)$.......	$-\dfrac{R}{Ls}E(s)$
$e_o(t)$........	$-\dfrac{R_f}{R_s}e(t)$........	$-RC\dfrac{de(t)}{dt}$	$-\dfrac{L}{R}\dfrac{de(t)}{dt}$......	$-\dfrac{1}{RC}\int e(t)dt$ ...	$-\dfrac{R}{L}\int e(t)dt$

Equation 3 can be used to determine the output current $i_o(t)$ for any input current $i_1(t)$ by use of the inverse transform. If the following inequalities are substantially satisfied

$$|A_c|>>\left|\frac{Z_f}{Z_o}\right|$$

$$\left|\frac{Z_f}{Z_p}A_c\right|>>\left|\left(1+\frac{Z_i}{Z_p}\right)\left(1+\frac{Z_f+Z_L}{Z_o}\right)+\frac{Z_f}{Z_p}\left(1+\frac{Z_L}{Z_o}-\frac{Z_{12}}{Z_o}\right)-\frac{Z_{12}}{Z_p}A_c\right|$$

then equation 3 can be approximated as equation 4

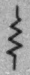
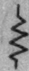
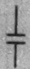
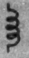
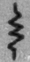
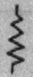
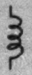
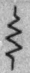
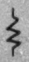
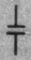
$$I_o(s)\approx-\frac{Z_p}{Z_f}I_i(s) \tag{4}$$

It should be noted that, to make equation 4 hold, the conditions on the amplifier (whose current gain is A_c) are that the input impedance should be low, the output impedance high, and $|Z_{12}|<<|Z_f|$. This latter condition is easy to satisfy in many practical cases. For example, a grounded emitter stage has a Z_{12} of about 30 ohms for a typical transistor. These conditions are as expected since they are the duals of the conditions for the voltage case. An amplifier consisting of junction transistors can come close to the ideal.

An application table similar to Table I can be constructed from equation 4. This table indicates the combinations of feedback impedances for use with the current operational amplifier to perform the operations of integration, multiplication by a scale factor, and differentiation. The table is seen to contain two possible integrators and two differentiators shown in such order that they are duals of the corresponding entries of Table I. If one is confined to the use of capacitors and resistors only for the critical feedback elements, then the circuits used for the current operational amplifiers are the duals of the inductance-resistance combinations of Table I.

Some physical feeling for the properties of this circuit can be had by considering the equations of input and output impedance which are derived in the second

Table II. Application Table for Current Operational Amplifier

Function	Amplifier	Differentiator a	Differentiator b	Integrator a	Integrator b
Z_f					
Z_p					
$I_o(s)$	$-\frac{R_p}{R_f}I(s)$	$-\frac{L}{R}sI(s)$	$-RCsI(s)$	$-\frac{R}{Ls}I(s)$	$-\frac{1}{RCs}I(s)$
$i_o(t)$	$-\frac{R_p}{R_f}i(t)$	$-\frac{L}{R}\frac{di(t)}{dt}$	$-RC\frac{di(t)}{dt}$	$-\frac{R}{L}\int i(t)dt$	$-\frac{1}{RC}\int i(t)dt$

and third parts of Appendix II. The input impedance (neglecting Z_{12}) at terminals $1'G$ in Fig. 2 is given by

$$(Z_{\text{in}})_{1'G}=Z_i+Z_f\frac{Z_L+Z_o(1-A_c)}{Z_f+Z_L+Z_o} \tag{5}$$

Let the quantities in equation 5 be taken as real. This would be required for the cases of integrator *b*, differentiator *a*, and the feedback amplifier or adder if only frequencies in the passband are considered. It is apparent that by making the current gain A_c a large negative number the input impedance can get quite large, a figure in the megohms being possible. Furthermore, if the load impedance R_L and feedback impedance Z_f are much smaller than Z_o, then we have the phenomenon of multiplication of the feedback impedance by the gain A_c of the system to give part of the impedance with respect to the terminals $1'G$. This is just the dual of the well-known Miller effect of the multiplication of the feedback admittance to give part of the input admittance of the conventional voltage circuit.

The output impedance equation, again neglecting Z_{12}, is

$$(Z_{\text{out}})_{0'G}=Z_o+Z_f-\frac{Z_fZ_o\left(A_c+\frac{Z_f}{Z_o}\right)}{Z_p+Z_i+Z_f} \tag{6}$$

For large real negative values of A_c it is found that high values of output impedance can be obtained even though moderate values of Z_o (the output impedance without feedback) are used. Thus, the output terminals of the feedback amplifier tend to behave like the terminals of a good current source.

The accuracy with which desired operations can be performed with the operational amplifier can theoretically be made as great as necessary since the difference between actual output and desired output approaches zero as A_c increases in magnitude. In practice, increasing A_c is always accompanied by diminishing bandwidth and increased stability problems.

One application of the operational amplifier not listed in Table II is that of addition. The adder is closely related to the first entry of Table II. An adder circuit is shown in block form in Fig. 3. The quantities to be added are currents, each of which can come from a current generator with internal resistance R_i. Subject to the inequalities cited previously, the ideal output current would be given by

$$i_o(t)=\frac{-R'R_p}{(R'+R_p)R_f}\left[i_1(t)+i_2(t)+\ldots i_i(t)+i_n(t)\right] \tag{7}$$

The coefficient of the sum of the input currents in equation 7 can be adjusted to give a suitable scale factor by varying R_p or R_f.

Much of what has been said about the problem of drift in transistor direct-voltage amplifiers applies to the case of the d-c operational amplifier. As in the voltage case, an odd number of grounded emitter stages should be included to make A_c a negative quantity. A suitable 2-stage combination might be a grounded base input stage followed by a grounded emitter. Such a combination would give a very low input impedance, a value of the order of 50 ohms being typical, and a current gain of over a hundred using high alpha junction transistors. The load impedance on the whole circuit should be as small as possible. Again, what has been said about cascading stages of voltage amplification holds, in general, for cascaded current amplifiers.

With all the advantages of ruggedness, reliability, and efficiency there are still several difficulties when using transistors in a large computing system. One of these has already been mentioned, i.e., the drift problem of d-c amplifiers must be used. Another difficulty in a large system using the current operational amplifier presented here is the necessity of noncommon power supplies. If several of these building blocks were connected together, this might be a serious problem. On the other hand, amplifiers of the type described are used in applications other than large-scale analogue computers. For example, a single operational amplifier type of circuit is often used in the generation of precision linear sweep waveforms and for driving function potentiometers in simple computing operations. Many of these jobs can be done with a-c coupling and with a single circuit of the type described. However, the possibilities of the extensive use of transistors, with the savings of power, size, and weight, are very attractive for large systems. The problems which must be solved before this is feasible do not seem insurmountable.

Appendix I. Derivation of Gain Equation of Voltage Operational Amplifier

Refer to Fig. 1, and write the transform node equations

$$\begin{aligned}&(E_1-E_i)Y_s+(E_1-E_o)Y_f+I_1=0\\&(E_o-E_1)Y_f+E_oY_L+I_2=0\end{aligned} \tag{8}$$

From 4-terminal network theory we have

$$\begin{aligned}&I_1=E_1Y_{11}+E_oY_{12}\\&I_2=E_1Y_{21}+E_oY_{22}\end{aligned} \tag{9}$$

Substituting equations 9 into 8 and solving for E_o in terms of E_i gives, after simplification

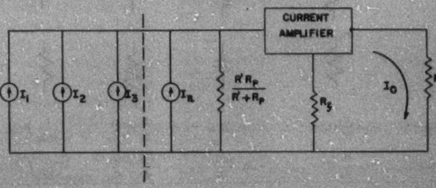


Fig. 3. Circuit arrangement for a current adder

$$E_o=\frac{E_1\left(\frac{Y_f}{Y_{22}}-\frac{Y_{21}}{Y_{22}}\right)}{\left(1+\frac{Y_{11}}{Y_s}\right)\left(1+\frac{Y_f+Y_L}{Y_{22}}\right)+\frac{Y_fY_L}{Y_sY_{22}}+\frac{Y_f}{Y_s}\left(1+\frac{Y_{21}}{Y_{22}}\right)+\frac{Y_{12}}{Y_s}\left(\frac{Y_f}{Y_{22}}-\frac{Y_{21}}{Y_{22}}\right)} \qquad (10)$$

Appendix II

Derivation of Gain Equation of Current Operational Amplifier

Refer to Fig. 2, and write the transform loop equations

$$\begin{aligned}&(I_1-I_t)Z_p+(I_1-I_o)Z_f+E_1=0\\&(I_o-I_1)Z_f+I_oZ_L-E_2=0\end{aligned} \qquad (11)$$

From 4-terminal network theory

$$E_1=I_1Z_{11}-I_oZ_{12}$$
$$E_2=I_1Z_{21}-I_oZ_{22}$$

Substituting these into equation 11 and solving for I_o

$$I_o=\frac{I_1\left(\frac{Z_f}{Z_{22}}+\frac{Z_{21}}{Z_{22}}\right)}{\left(1+\frac{Z_{11}}{Z_p}\right)\left(1+\frac{Z_f+Z_L}{Z_{22}}\right)+\frac{Z_fZ_L}{Z_pZ_{22}}+\frac{Z_f}{Z_p}\left(1-\frac{Z_{21}}{Z_{22}}\right)-\frac{Z_{12}}{Z_p}\left(\frac{Z_f}{Z_{22}}+\frac{Z_{21}}{Z_{22}}\right)} \qquad (12)$$

Derivation of Input Impedance Between Terminals 1′G

Consider a potential source applied between terminals 1′G. The ratio of this voltage to I_1 is the following input impedance

$$(Z_{in})_{1'G}=\frac{I_1Z_{11}-I_oZ_{12}+(I_1-I_o)Z_f}{I_1}$$

where

$$\frac{I_o}{I_1}=\frac{Z_{21}+Z_f}{Z_f+Z_L+Z_{22}}$$

After substitution and simplification

$$(Z_{in})_{1'G}=Z_{11}+\frac{Z_f\left[Z_L+Z_{22}-Z_{12}-Z_{21}-\frac{Z_{12}Z_{21}}{Z_f}\right]}{Z_f+Z_L+Z_{22}} \qquad (13)$$

Note that since $Z_{11}=Z_i$, $Z_{22}=Z_o$, $Z_{21}=Z_oA_c$ and if Z_{12} is neglected equation 13 reduces to equation 5.

Derivation of Output Impedance

A current $-I_o$ is applied to terminals $0'G$. The voltage between $0'G$ is

$$V_{o'G}=I_1Z_{21}-I_oZ_{22}+(I_1-I_o)Z_f$$

$$(Z_{out})_{o'G}=\frac{V_{o'G}}{-I_o}=Z_{22}+Z_f-\frac{I_1}{I_o}(Z_{21}+Z_f)$$

$$\frac{I_1}{I_o}=\frac{Z_{12}+Z_f}{Z_p+Z_f+Z_{11}}$$

Hence

$$(Z_{out})_{o'G}=Z_{22}+Z_f-\frac{(Z_{12}+Z_f)(Z_{21}+Z_f)}{Z_p+Z_f+Z_{11}} \qquad (14)$$

If Z_{12} is neglected and it is recalled that $Z_{21}=Z_oA_c$, $Z_{22}=Z_o$, $Z_{11}=Z_i$, then equation 14 reduces to equation 6.

References

1. Electronic Analogue Computers (book), **G. A. Korn, T. M. Korn.** McGraw-Hill Book Company, Inc., New York, N. Y., 1952.

2. Analysis of Problems in Dynamics By Electronic Circuits, **Ragazzini, Randall, Russell.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 35, May 1947.

3. Transformation Calculus and Electrical Transients (book), **S. Goldman.** Prentice Hall, Inc., New York, N. Y., 1949.

4. Frequency Variations of Current Amplification Factor for Junction Transistors, **R. L. Pritchard.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 40, Nov. 1952.

5. D-C Amplifier Employing Junction-Type Transistors, **Edward Keonjian.** *Electrical Engineering*, vol. 72, no. 11, Nov. 1953, pp. 961–64.

6. Principles of Transistor Circuits (book), **R. F. Shea, et al.** John Wiley and Sons, Inc., New York, N. Y., 1953.

◆

No Discussion

A Simplified Standard Cell Comparator

J. H. MILLER
FELLOW AIEE

Synopsis: For use in comparing working standard cells with the potential from a "normal" or saturated cell, presumably one of a group maintained as a voltage reference, a simplified comparator is presented. Complete circuit details are given, as well as a description of the necessary auxiliary items.

THE use of the saturated or normal form of Weston standard cell, with its long life and well-maintained voltage, appears to be growing. A decade ago relatively few banks of saturated cells existed outside the National Bureau of Standards. Today, however, numerous industries and college laboratories are maintaining groups of saturated cells for the immediate availability of a high-accuracy voltage reference.

Saturated cells, while maintaining their voltage well against time in years, must be maintained in a temperature-controlled chamber since the change in voltage with temperature is about 60 microvolts (μv) per degree centigrade. Through the use of the Wolff formula for temperature correction and occasional cross checks at the Bureau of Standards, an independent group of saturated cells can readily maintain the volt to within 5 μv. The requirement for constant temperature, however, precludes the use of such cells in completely portable service and for general use the unsaturated cell with its low temperature coefficient is more suitable. For standardizing the unsaturated cells in terms of the values of a normal saturated bank, a simple comparator of some sort is desirable.

Definition of a Comparator

Stated very simply, in checking a working standard cell against a reference standard, the two cells are connected in opposition, in series with a galvanometer and an auxiliary source of a few μv. Adjustments are made in the auxiliary source until the galvanometer indicates a null balance. The working cell potential then equals the potential of the reference cell plus, algebraically, that of the auxiliary potential. The comparator is simply a means of adjusting the auxiliary potential, indicating its value in μv in some manner and possibly with an arrangement whereby it can be added simply, accurately, and with proper sign, to the value of the reference cell.

The comparator designed by H. B. Brooks[1] is a highly developed form in which readings can be made to 0.1 μv. Used for research and for international comparison of voltage standards, it is necessarily of high accuracy. It includes some 70 adjusted resistors and a rather elaborate calculating system.

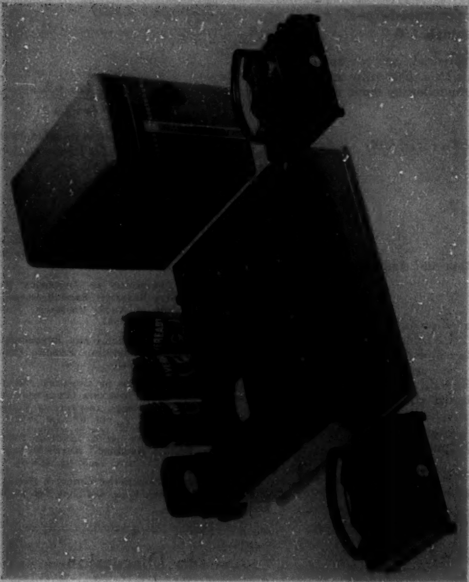
Statement of Requirements

A useful comparator of simpler scope might have a sensitivity level of 1 μv

---

Paper 54-208, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted February 15, 1954; made available for printing April 7, 1954.

J. H. Miller is with the Weston Electrical Instrument Corporation, Newark, N. J.



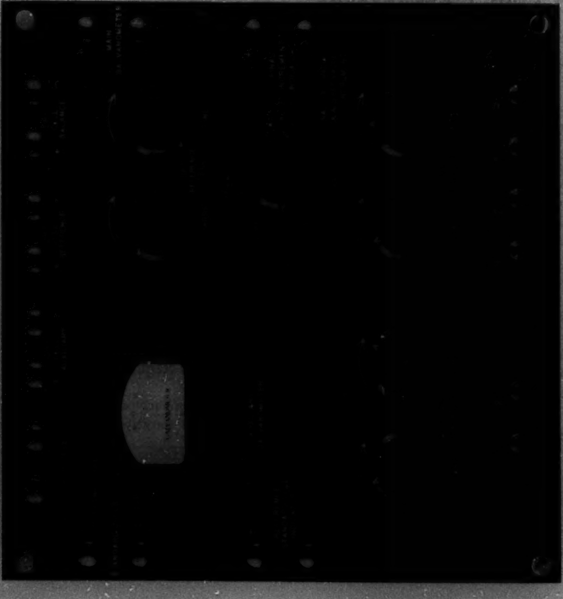


Fig. 1 (left). Top panel of standard cell comparator

Fig. 2 (above). Comparator with associated milliammeters, dry cells, main galvanometer, and auxiliary standard cell

with 5 μv per division on the indicating instrument and an over-all limit of error of 5 μv. On the assumption that the standard cell being checked will come within the extreme limits of 1.017000 and 1.020500 volts for a span of 3,500 μv it is possible to get along with seven coarse steps each covering 500 μv, and the associated instrument covering the fine subdivision of each 500-μv step. If it is assumed that the saturated cell bank is being maintained at least between 28.2 and 36.6 degrees centigrade, then, according to the Wolff formula, the reference voltage span will be 1.017750 and 1.018250 amounting to, again, 500 μv and centered at 1.018000 volts.

Assembling these design parameters, the comparator illustrated in Fig. 1 was built. It is assembled with the necessary auxiliary items in Fig. 2, including the two milliammeters, the auxiliary standard cell, three dry cells, and the main galvanometer of high sensitivity; the interconnecting wiring has been omitted for clarity. Fig. 3 is a schematic diagram including all of the auxiliary items. Open circles represent binding posts, and true voltage values are given at the switch points.

The circuit of the reference cell and the cell under test, X, is closed to the high-sensitivity galvanometer G through the resistance from A to B which actually takes the form of a few feet of no. 24 gauge manganin wire wound on a skeleton bakelite form to give about 1/4 ohm per turn. A tap is made on each turn, the assembly is varnished and baked, and the sections marked 0.25 are adjusted to better than 0.001 per cent or 1/2 μv in 500 when passing 2 milliamperes (ma). Sections marked *na* are not adjusted.

The left-hand model *1* instrument has a range of 1-0-1 ma. Curent covering this span, controlled by the associated rheostats, produces a drop in the 0.25-ohm resistor of from 250 μv aiding, through zero, to 250 μv, opposing, the reference cell. The instrument scale is marked, however, 1.017750 to 1.018250 volts, with 1.018000 volts at the electrical zero point. The current is adjusted so that the instrument indicates the actual reference cell potential at the then existing temperature; the drop of this measured current through the 0.25-ohm resistors then, added or subtracted from the reference cell, adjusts the net voltage to 1.018000 volts.

Current through the center 0.25-ohm resistors, as selected by the rotary switch, is maintained at exactly 2 ma by balancing the drop of this current through the 509.85-ohm resistor against the potential

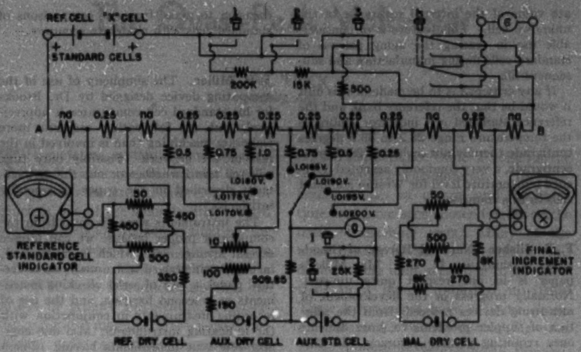


Fig. 3. Schematic diagram of comparator with all auxiliary items

of an auxiliary standard cell through a simple galvanometer *g*, exactly as in the Brooks design. The resistors in the switch point leads, 0.5, 0.75, etc., serve to maintain constant resistance in the battery circuit on any switch position, thereby obviating readjustment of the 2-ma current as the range is changed.

Selection of the switch position adds or subtracts voltage in 500-μv steps to bring the reference voltage to the values indicated at the switch points on the diagram, and as necessary for a coarse balance on the main galvanometer *G*.

The final adjustment to balance is by means of the network at the right. When balance has been achieved, the current in this network producing the final μv value is measured in the right-hand instrument which, in turn, is marked 0–500 μv in black and coded as scale *A*, and 500–1,000 μv in red and coded as scale *B*. This simplifies the use of the correct set of numbers since at the switch point for 1.0180 volts the white-filled engraving reads $1.018+A$, whereas after the point for 1.0185 volts the red-filled engraving reads $1.018+B$; see Fig. 1. The other points are similarly marked.

The main galvanometer has a resistance of 25 ohms and its external critical damping resistance value of 400 ohms roughly matches the resistance of the two cells in series. One scale division deflection, 0.005 microampere, would be produced by 2 μv through 400 ohms, thus indicating the degree of resolution.

Accuracy of Adjustment

A brief analysis of the circuit will show that only the marked resistors in the circuit from *A* to *B* and the 509.85-ohm resistor need be adjusted accurately. They were held to 0.1 per cent. All of the other resistors serve adequately if within a few per cent of the marked value including the control rheostats. On the other hand, the values in the dry-cell circuits were selected after considerable study and actual use with the instrument to give the most satisfactory manipulation under all conditions. In each case the upper rheostat is the fine adjustment.

The device was built with full knowledge that Dr. Brooks designed an elaborate temperature-compensated galvanometer key, that he furnished a compensator for any residual thermal voltage, and that he was much concerned about various random residual potentials. It was expected that possibly copper binding posts would be required, along with other special compensation means. However, practical use in a laboratory and checks of many cells as manufactured appear to indicate that if accuracy to within a few μv is considered adequate no further compensation is needed. To be sure, the apparatus is left connected all of the time and, as Dr. Brooks points out, the battery drain is no more than equivalent to shelf life. Thus the apparatus reaches temperature equilibrium and is touched in normal use only on the control knobs and the galvanometer keys which are made of bakelite.

In presenting this particular form of comparator and its circuit, full credit must be given to Dr. Brooks for his original paper from which this design is abstracted. Nevertheless since this simpler form is the result of considerable study, it is presented as a useful variant for the record.

Reference

1. **The Standard Cell Comparator, a Specialized Potentiometer, H. B. Brooks.** *Research Paper No. 586,* Journal of Research, National Bureau of Standards, Washington, D. C., vol. 11, Aug. 1933, p. 211.

◆

Discussion

H. B. Brooks (Consulting Specialist, Washington, D. C.): The comparator described by Mr. Miller is a very good solution to the problem of checking standard cells with an accuracy adequate for the needs of the maker of such cells, and for testing institutions which are not obliged to push the accuracy to the extreme limits for which national standardizing laboratories must strive. To a large extent it avoids the special design features which were found necessary in the standard cell comparator as used at the National Bureau of Standards (NBS). It is one more illustration of the usefulness, for certain cases, of Poggendorff's "second method," which remained unnoticed for 70 years until its possibilities were brought out by Dr. Lindeck of the German Reichsanstalt. The Lindeck element has undeveloped possibilities as an "element" in the design of potentiometers for various purposes. Among its advantages are the ease with which it may be given more than one range and the facility with which it may be made extremely free from error caused by thermal electromotive forces.

As illustrations of the versatility, convenience, and freedom from parasitic electromotive forces (emf) of the Lindeck element, I might cite two further applications that were developed at the National Bureau of Standards. (1) In a low-range potentiometer constructed for thermocouple work,[1] a Lindeck element was incorporated having six ranges with full-scale values from 40 μv to 2 millivolts. With suitable thermal shielding the parasitic emf in the measuring circuit was well below the level of 0.1 μv. (2) A Lindeck element having three ranges with full-scale values from 100 μv to 10 millivolts was used as the final element in a wide-range potentiometer designed for general voltage measurements and described in an unpublished paper by Brooks, Harris, and Schroyer. The parasitic emf introduced into the measuring circuit by this element was below the level of 0.01 μv.

The author has not made use of the computing device which I incorporated in the NBS standard-cell comparator as he considers it "rather elaborate." Actually this device consists essentially of a mask and a second set of engraved numbers on the step dial of the potentiometer. It is true that this would require additional shop work, and it may well be rather elaborate from this viewpoint. However, in use it is quite simple, requiring no additional manipulations in the balancing procedure; and it has the very decided advantage, from the user's point of view, that it algebraically adds the measured difference of potential to the emf of the reference cell and thus gives directly the emf of the cell under test. This automatically prevents reading errors that might occasionally occur if the observer were required to perform the necessary addition or subtraction as a part of the emf determination. It therefore appears to me that some further consideration might be given to the incorporation of such a device in any commercial realization of Mr. Miller's comparator.

I should like to hear Mr. Miller's reasons for using separate, model *I* instruments as his "associated milliammeters." They are actually indispensable components of his comparator. In special potentiometers as designed at the NBS we have found that suitably modified, built-in fan-shaped switchboard instruments have given good service, and (according to our taste) have given an air of unity to the instrument. The desirable modifications include: knife-edge pointer, parallax mirror, fine division lines as in standard portable instruments, and readily accessible zero adjuster.

REFERENCE

1. A MULTIRANGE POTENTIOMETER AND ITS APPLICATION TO THE MEASUREMENT OF SMALL TEMPERATURE DIFFERENCES, H. B. **Brooks**, A. W. **Spinks**. *Journal of Research*, National Bureau of Standards, Washington, D. C., vol. 9, 1932, p. 781.

R. C. Langford (Weston Electrical Instrument Corporation, Newark, N. J.): The variant of the Brooks comparator by Mr. Miller has been in use for routine checking and analysis of several hundred Weston standard cells in both manufacturing and laboratory areas. The ease of reading, together with the lack of ambiguity in not requiring to know whether to add or subtract the voltage divergence from the reference cell, adds considerable merit to its utility and reliability.

Although the design level of accuracy is some 5 μv, actual errors of this order of magnitude may only be reached when the divergence of the voltage of the unknown cell from the reference cell is large. This feature is of importance since the paper does not specifically point out that normal as well as unsaturated cells may be checked. Thus the voltage, e.g., of a normal cell, newly arrived back from the NBS, can be checked against the reference normal cell of the comparator to a considerably higher accuracy than the 5 μv claimed.

A feature of the use of the comparator is that the difference between the voltage of the unknown cell with and without a 1-megohm resistor across it gives an immediate value of the internal resistance of the unknown cell. This feature is of considerable use in tracing the aging process of standard cells during manufacture and subsequent life.

If any criticism is to be made of this unit it would be that the reference bank and the reference standard cell indicator embraces the temperature range 28.2 to 36.6 degrees centigrade thereby not including 28 degrees centigrade, a very commonly accepted reference temperature level for Weston standard cells in this country.

F. B. Silsbee (National Bureau of Standards, Washington, D. C.): Mr. Miller's paper is in one respect quite unusual. Normally progress in the development of measuring devices proceeds from the solution of simpler problems to more difficult ones requiring greater accuracy. In this case, on the other hand, the more difficult problem of comparing standard cells quickly and with an accuracy of 0.1 μv was solved in 1933 by Dr. H. B. Brooks. Now, however, a new demand has arisen for measurements similar in kind but of less exacting requirements, and Mr. Miller has come up with a correspondingly simpler and, I presume, less expensive and quite adequate solution.

The Brooks comparator was designed to provide the ultimate accuracy needed for such specialized tasks as the intercomparison of the cells in the primary group by which the national standard of emf is maintained and for determining the small differences among the realizations of the volt in the national laboratories of different countries. It has proved highly satisfactory for this purpose. During the last 14 years its calibration has not changed more than can be compensated by an adjustment to the auxiliary standard cell dial equivalent to 0.6 μv in the unknown cell.

The Miller design reduces the number of adjusted resistors and contact studs by a factor of 5 and does not use a thermal shield or a special thermo-free galvanometer key. On the other hand its accuracy is less by a factor of 5. Its range of 2,500 μv is ample for present conditions while the Brooks design covered 3,400 μv to allow for the change which occurred January 1, 1948, from the International Volt to the Absolute Volt. Brooks had anticipated this change and planned accordingly. The use by Miller of a red and a black scale on one milliammeter should be adequate to minimize false readings although it is not as completely foolproof as is Brooks' very simple shield which allows the user to see and read only the value of the two cells being compared.

This new instrument should prove very useful in the many laboratories which have occasion to make frequent comparisons of standard cells with moderate accuracy.

J. H. Miller: The simplicity of use of the computing device designed by Dr. Brooks for his standard cell comparator is appreciated, but its fabrication requires more elaborate machining than is involved in the rest of the network. Possibly only time will tell as to whether or not the arrangement described in the paper is generally satisfactory.

The instruments were divorced from the comparator proper because we can foresee several arrangements which might require, in turn, switchboard instruments on a vertical panel, a pair of series checking instruments at a second location, and the use of recording instruments in conjunction with the indicating instruments; also one engineer felt that instruments having 12-inch scales might be advisable from the viewpoint of easier readability. Further, having the instruments separate allows for them to be adjusted or even serviced if such is necessary and perhaps replaced by other instruments. Perhaps this is all in the point of view and the matter of the inclusion of the instruments or not in a common panel is possibly beside the point being made here on the simplified circuit diagram and manipulation.

The limitation of the referenced standard cell indicator to potentials embracing the temperature range of 28.2 to 36.3 degrees centigrade, was purely to simplify the adjustment of the instrument to 250 μv each side of 1.018000 volts. In the Weston Laboratories 33 degrees centigrade is used as the temperature at which the reference cells are held and which eliminates the need of cooling. If it is desired that the reference standard cell indicator cover a potential corresponding to 28 degrees centigrade, as suggested by Dr. Langford, the center electrical zero could be displaced to one side a bit. Actually the value of 1.018 volts could be placed at either end of the instrument scale and considered as the high limit or the low limit without in any way invalidating the adjustments or the markings on the comparator proper. Perhaps this is another reason for the separate instruments.

Placing a 1-megohm resistor across the cell under test is, of course, an excellent method of determining its internal resistance. Although such a resistor was not a part of the comparator described here, it is possible to add it to the structure with the addition of another terminal and a few minor items.

In closing, the writer is in thorough agreement with the summary given in the discussion by Dr. Silsbee.

The Quasi-Peak Voltmeter

C. W. FRICK
MEMBER AIEE

THE metering system commonly used in radio noise meters includes a special type of voltmeter, the development of which is one of the steps leading toward practical instruments for evaluating the influence of electric power apparatus on radio reception. Early developments in this field were introduced through the efforts of groups representing the electric-apparatus manufacturers, power companies, radio apparatus manufacturers and services. Radio noise meters and test methods have been gradually improved as a result of experience, and progress reported from time to time. General features of radio noise meters and test methods were discussed in 1943.[1] The over-all performance of a representative type of noise meter, using square wave input voltage to simulate noise, was analyzed in 1945.[2] Several meter circuits of the type applied to radio noise meters are analyzed in this paper.

The need for attention to the metering system, which is the final stage of the noise meter circuit, became evident when it was attempted to measure certain types of noise. The early instrument was substantially a receiver with a voltmeter to measure the output. Difficulty was experienced in measuring noise characterized by spaced pulses such as are produced by a gas engine ignition system. This experience led to the development of a weighting circuit to actuate the voltmeter or indicating instrument. The weighting circuit is a capacitance-resistance network with charge time of the order of 1 to 10 milliseconds (ms) and discharge time of the order of 100 to 1,000 ms. The charge time and the discharge time are the time constants as usually defined for a capacitor with resistance. Readings taken with this type of weighting network are often called quasi-peak values. In this paper the indicating instrument in combination with the circuit which constitutes the output stage of the noise meter is called a quasi-peak voltmeter.

The object of the metering system is to give readings which are reasonably significant to the nuisance value of noise for the types usually encountered. The particular values of charge and discharge time used have been chosen largely on the basis of listening tests. Many of the noise meters in use at present in this country were built according to specifications in a report published in 1940.[3] These instruments have a 10-ms charging time and a 600-ms discharging time. Noise meters developed more recently, some of which are for frequencies above or below the range of the 1940 report, are designed in accordance with proposed specifications[4,5] issued by the American Standards Association (ASA) for trial use. The ASA proposal is intended to extend and improve the 1940 specifications. Among other modifications the specified charge time is 1.0 ms instead of 10. The value of 1.0 ms for charging is used in the specification for an instrument to measure radio noise adopted by the International Special Committee on Radio Interference.[6] The time for discharging was originally 160 ms but the value has been changed to 500 ms. In general, the noise meters recently developed in this country and abroad have a 1.0-ms charge time and a 600- (or 500-) ms discharge time.

The correlation between the nuisance value of noise and the characteristics of noise meters to be specified is still subject to some uncertainty. Further investigation might be undertaken in stages. For example, in one part of the investigation the factors significant to the nuisance value of common types of radio noise disturbances might be evaluated with the aid of improved testing equipment and techniques now available. In another part of the investigaton the readings which representative noise disturbances produce on different instruments might be determined. A simplified method of analyzing instrument performance has been described,[2] and applied to square-wave noise sources. To include all noise sources further work will have to be done so that irregular or random noises can be treated. Meanwhile, the results obtained with noise consisting of uniformly repeated wave fronts or pulses are useful for some purposes, such as the correlation of readings taken with different instruments. This paper presents an analysis of the performance of the output stage of the noise meter on the basis of uniform pulses. Data are included which enable calculations to be made for conditions to which the analysis is applicable.

Explanation of Term Quasi-Peak

The term quasi-peak as applied to voltmeters in this discussion indicates that the instrument has characteristics such as those of a peak-reading voltmeter for pulsating voltage but does not register the true peak value.

Theoretically a d-c voltmeter connected in parallel with a small capacitor across a pulsating voltage would register the peak voltage, but this ideal characteristic is not realized in practice because it requires a finite time to charge the capacitor to a given voltage and because the capacitor loses a small but appreciable portion of its charge between peaks.

Although the quasi-peak method is sometimes thought of as a modified peak method, it actually evaluates a pulsating voltage in terms of several characteristics, one of which is the maximum or peak of the pulse voltage. Other characteristics are the frequency of occurrence of the peaks usually called the pulse repetition rate, the length of time the peak lasts, and the shape of the pulses which make up the pulsating voltage. The voltmeter and associated weighting circuit can be designed to evaluate a pulse voltage in a number of different ways. For example, the instrument can be designed so that the reading increases with pulse repetition rate ultimately approaching the true peak value. This type of evaluation is used in the radio noise meter because the effect of noise pulses on the ear increases with pulse repetition rate.

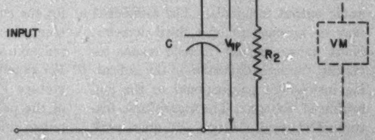


Fig. 1. Basic circuit of quasi-peak voltmeter

Paper 54-204, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 1, 1954; made available for printing March 30, 1954.

C. W. FRICK is with the General Electric Company, Schenectady, N. Y.

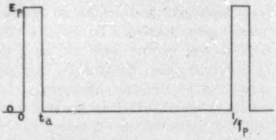
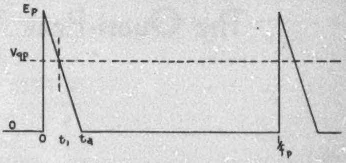


Fig. 2 (left). Rectangular pulse

Fig. 3 (right). Saw-tooth pulse



Basic Circuit

A circuit which gives the response characteristics just described is shown in Fig. 1. A voltage pulse of the proper polarity applied to the input terminals produces a charge on the capacitor C. The time required to charge this capacitor is determined by the product R_1C where R_1 is the resistance through which the capacitor is charged. The effect of the shunting resistance R_2 on the charge time is neglected because R_2 must be at least 10 times the value of R_1 to obtain the characteristic usually desired. When the applied voltage is removed, the capacitor is discharged through the resistance R_2. The time required to discharge the capacitor is determined by the product R_2C. The rectifier A precludes the possibility of the capacitor discharging through the input circuit.

When voltage pulses of equal magnitude and regular recurrence are applied to the input, the capacitor C is charged to a voltage such that the voltage increment for each pulse is equal to the voltage decrement between pulses, as explained in the analysis which follows. When this steady-state condition is reached, the average capacitor voltage is the quasi-peak voltage referred to in the preceding section. The capacitor voltage can be measured by a separate voltmeter as indicated at VM in Fig. 1 or, by suitable choice of R_1, R_2, and C, the resistance R_2 may be the series resistance of a conventional d-c voltmeter. If a separate voltmeter is used, it should preferably be of the electronic type wherein the voltage is impressed on the grid of a vacuum tube.

Most radio noise meters employ the superheterodyne receiver circuit in which case the rectifier A in Fig. 1 is the second detector and the resistance R_1 includes the resistance of the detector tube. An alternative arrangement, not generally used in the United States, is to connect the second detector to an audio stage and connect the voltmeter of Fig. 1 to the audio output terminals. The rectifier A may be an element of a third detector. Noise meters usually include means for making the scale deflection of the indicating instrument proportional to the logarithm of voltage. The logarithmic feature affects peak voltages and quasi-peak voltages alike and therefore the theory developed applies either to the linear scale or to the logarithmic scale when scale readings or voltages at the noise meter input are used.

Pulse Response

The term pulse response is used in this discussion to designate the ratio of quasi-peak voltage to peak voltage when voltage pulses are applied at a uniform repetition rate and a steady-state condition is reached. This ratio is a function of the wave shape and time duration of the pulse, the pulse repetition rate, and the ratio of discharge time to charge time. In the sections which follow, the pulse response will be determined for several types of pulses.

Rectangular Pulses

One type of pulse which might be applied to the circuit of Fig. 1 is a rectangular pulse such as shown in Fig. 2. If the polarity is in the right direction to go through the rectifier A, a charge is accumulated on the capacitor C. The magnitude of the pulse is represented by E_p, the peak voltage. The number of pulses per second if f_p and therefore the time of one pulse cycle is $1/f_p$ as indicated in Fig. 2. The pulse duration t_a is short compared with the time of 1 pulse cycle. The analysis given here is applicable when the ratio of these quantities is less than 0.2. The ratios encountered in radio noise measurements are usually much less than this value.

Starting with zero voltage, the first pulse charges the capacitor to voltage V_1 where

$$V_1 = E_p\left(1-\epsilon^{-\frac{t_a}{R_1C}}\right) \tag{1}$$

The other quantities are indicated in Fig. 2.

When the pulse width t_a is small compared with the charge time constant R_1C for the circuit of Fig. 1, which is usually true, the voltage produced by the first pulse is a small fraction of the peak voltage. For example, if t_a is one-tenth of R_1C, the voltage V_1 is of the order of 10 per cent of the peak voltage. For a sequence of pulses and in the absence of discharge the capacitor voltage would gradually increase to the full peak value.

During the interval between pulses the capacitor is discharged exponentially at a rate determined by R_2C, the discharge time constant for the circuit of Fig. 1. Since the analysis applies to relatively short pulses, the interval for discharge is approximately $1/f_p$ (see Fig. 2). Usually the value of R_2C is at least ten times the interval between pulses. Under these conditions the decrease in voltage during the discharge period is approximately

$$\Delta V = \frac{V}{f_pR_2C} \tag{2}$$

where $V=$ the voltage on the capacitor. For the usual discharge rate ΔV is less than 10 per cent of V.

When the capacitor is already charged to voltage V, the applied voltage during a pulse is $E_p - V$. Successive pulses, tend to reduce this difference, which causes the capacitor voltage to approach a steady-state value V_{qp}, the quasi-peak voltage. This voltage can be determined by equating the increase in capacitor voltage for each pulse to the decrease between pulses. The voltage increment is approximately

$$\Delta V' = \frac{(E_p - V_{qp})t_a}{R_1C} \tag{3}$$

The voltage decrement is given by equation 2 with V_{qp} in place of V. Combining these equations we obtain

$$\frac{E_p}{V_{qp}} = \frac{1}{t_af_p\rho}+1 \tag{4}$$

where $\rho = R_2C/R_1C$, the ratio of discharge time to charge time.

This equation may be written

$$P = \frac{t_af_p\rho}{t_af_p\rho+1} \tag{5}$$

where $P = V_{qp}/E_p$, the pulse response ratio.

The equation just developed can be used to determine the quasi-peak reading corresponding to a defined rectangular pulse when the constants for the circuit of Fig. 1 are known. Also, the equations may be used to check either the charge time or the discharge time of particular noise meters in the following manner. A sequence of pulses of measured width and repetition rate is impressed on the circuit by the procedure described in the example. The quasi-peak voltage and the peak voltage are read with the noise meter. The ratio of these readings is the value of P. Equation 5 is solved for ρ which gives

$$\rho = \frac{P}{t_a f_p (1-P)} \tag{6}$$

The value of ρ is the ratio of discharge time to charge time, and therefore if one of these quantities is known the other is determined. The discharge time constant R_2C is usually easy to obtain from resistance and capacitance measurements. The charge time constant R_1C is not so easy to obtain because R_1 usually includes the equivalent resistance of associated vacuum tubes. This test provides a means of checking the charge time under operating conditions. The pulse should be so chosen that the value of P is not too close to 1.0.

Saw-tooth Pulses

In the preceding section the pulses were considered as rectangular or a close approximation to the rectangular shape. Another pulse shape which is of at least theoretical interest is the saw-tooth pulse illustrated in Fig. 3. The pulse voltage is initially E_p and is reduced to zero at a uniform rate in time t_a.

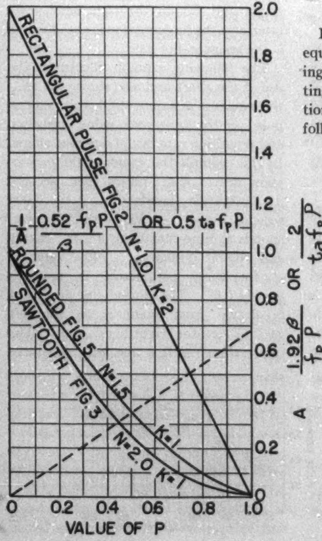


Fig. 4 (left). Graphical method for determining pulse response

In this case, the charging period is less than t_a since it ends when the pulse voltage has dropped to the voltage on the capacitor. It may be seen by referring to Fig. 3 that after the capacitor voltage has reached a steady-state value V_{qp} the charging period ends when

$$E_p\left(\frac{t_a - t_1}{t_a}\right) = V_{qp}$$

Solving for t_1 and replacing V_{qp}/E_p by P we have

$$t_1 = t_a(1-P) \tag{7}$$

The increment of voltage for each pulse, after the steady-state condition is reached is obtained in the following manner. The current at time t is equal to the net voltage divided by R_1. The net voltage is the difference between the pulse voltage at time t and the voltage on the capacitor V_{qp}. Substituting values for instantaneous voltage and capacitor voltage V_{qp}, the current at time t is

$$i = \frac{E_p}{R_1 t_a}(t_1 - t)$$

Integrating idt from zero to t_1 and dividing by C gives the increment of voltage which is

$$\Delta V' = \frac{E_p t_1^2}{2R_1 C t_a}$$

Substituting the value of t_1 from equation 7, we have

$$\Delta V' = \frac{0.5 E_p t_a (1-P)^2}{R_1 C} \tag{8}$$

For the steady-state condition $\Delta V'$ is equal to ΔV, the voltage decrement during discharge given by equation 2. Putting $V = V_{qp}$ for the steady-state condition, and substituting P for V_{qp}/E_p the following equation is obtained

$$\frac{P}{(1-P)^2} = 0.5 f_p \rho t_a \tag{9}$$

This equation can be solved for P the pulse response by graphical or other means. A graphical method which will be explained in another section, is shown in Fig. 4.

Rounded Pulses (Radio Noise)

Radio noise characterized by regularly repeated steep-wave fronts as in square waves or very short pulses such as 0.1-microsecond (μs) pulses, results in pulses at the noise meter output of the shape illustrated in Fig. 5. This curve is the envelope of the rectified output of the receiver section of the noise meter. It is given by the following equation derived by the author in 1945[2]

$$e = E_p \frac{\sin \pi\beta t}{\pi\beta t} \tag{10}$$

In this equation β represents the bandwidth of the radio noise meter in cycles. Frequencies within this band are passed by the receiver section of the noise meter, and other frequencies are cut off. The pulse width corresponding to t_a in Figs. 2 and 3 is a function of the bandwidth and is equal to $1/\beta$.

It may be mentioned that the analysis from which equation 10 was derived was based on square waves. The square wave was represented by the well-known Fourier series. A similar analysis can be made starting from the series for some other wave. For example, the Fourier series for 0.1-μs pulses, which may be obtained from information given in text books and handbooks, gives the same result as equation 10.

When a succession of pulses similiar to Fig. 5 is applied to the circuit of Fig 1, the charging period ends when the voltage e given by equation 10 is equal to the voltage on the capacitor. In the steady-state condition, the capacitor voltage is approximately equal to the quasi-peak voltage V_{qp}.

Following the procedure used for the saw-tooth pulse, the time t_1, at which the charging period ends for the steady-state condition, may be found from the equation

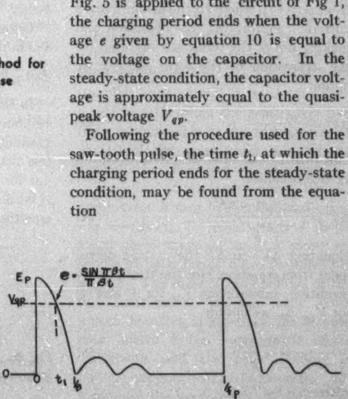


Fig. 5 (right). Rounded pulses produced by short pulses or steep wave fronts impressed on a radio noise meter

Table I. Table of Values of Pulse Response P for Radio Noise Meters

	0	1	2	3	4	5	6	7	8	9
$\frac{0.52 f_p \rho}{\beta}$										
0	0	0.01	0.02	0.03	0.04	0.05	0.06	0.06	0.07	0.08
0.1	0.09	0.09	0.10	0.11	0.12	0.12	0.13	0.14	0.14	0.15
0.2	0.16	0.16	0.17	0.17	0.18	0.18	0.19	0.20	0.20	0.21
0.3	0.21	0.22	0.22	0.23	0.23	0.23	0.24	0.24	0.25	0.25
0.4	0.26	0.26	0.27	0.27	0.27	0.28	0.28	0.28	0.29	0.29
0.5	0.30	0.30	0.30	0.31	0.31	0.31	0.32	0.32	0.32	0.33
0.6	0.33	0.33	0.34	0.34	0.34	0.35	0.35	0.35	0.35	0.36
0.7	0.36	0.36	0.36	0.37	0.37	0.37	0.38	0.38	0.38	0.38
0.8	0.39	0.39	0.39	0.39	0.40	0.40	0.40	0.40	0.40	0.41
0.9	0.41	0.41	0.41	0.42	0.42	0.42	0.42	0.42	0.43	0.43
1.0	0.43	0.43	0.43	0.44	0.44	0.44	0.44	0.44	0.45	0.45
$\frac{1.92\beta}{f_p \rho}$										
1.0	0.43	0.43	0.43	0.42	0.42	0.42	0.42	0.42	0.42	0.41
0.9	0.45	0.45	0.45	0.45	0.44	0.44	0.44	0.44	0.43	0.43
0.8	0.48	0.47	0.47	0.47	0.47	0.46	0.46	0.46	0.46	0.45
0.7	0.50	0.50	0.50	0.49	0.49	0.49	0.49	0.48	0.48	0.48
0.6	0.53	0.53	0.53	0.52	0.52	0.52	0.51	0.51	0.51	0.51
0.5	0.57	0.56	0.56	0.56	0.55	0.55	0.55	0.54	0.54	0.54
0.4	0.61	0.61	0.60	0.60	0.59	0.59	0.58	0.58	0.58	0.57
0.3	0.66	0.65	0.65	0.64	0.64	0.63	0.63	0.62	0.62	0.61
0.2	0.72	0.72	0.71	0.70	0.70	0.69	0.68	0.68	0.67	0.67
0.1	0.81	0.80	0.79	0.78	0.77	0.76	0.76	0.75	0.74	0.73
0	1.00	0.94	0.92	0.91	0.89	0.88	0.86	0.85	0.84	0.83

$P = \frac{\text{quasi-peak}}{\text{peak}}$

f_p = pulse repetition rate, pulses per second

β = bandwidth, cycles

$\rho = \frac{\text{discharge time}}{\text{charge time}}$

$$P = \frac{\sin \pi\beta t_1}{\pi\beta t_1} \qquad (11)$$

where $P = V_{qp}/E_p$.

The right-hand number of equation 11 can be replaced by a series. In this case the first two terms of the series give an approximation which is good enough for the purpose of this analysis. The equation then becomes

$$P = 1 - \frac{(\pi\beta t_1)^2}{6}$$

Solving for t_1 we obtain

$$t_1 = 2.45\sqrt{1-P}/\pi\beta \qquad (12)$$

The increment of capacitor voltage $\Delta V'$ for each pulse shown in Fig. 5 is calculated as before by integrating idt from zero to t_1 where i is the instantaneous current. Performing this operation and making use of the series already used for the expression of the form $\sin x/x$ we obtain

$$\Delta V' = \frac{E_p(\pi\beta t_1)^2}{9\pi\beta R_1 C} \qquad (13)$$

Substituting the value of t_1 from equation 12 this becomes

$$\Delta V' = \frac{0.52 E_p (1-P)^{3/2}}{\beta R_1 C} \qquad (14)$$

Equating $\Delta V'$ to ΔV the voltage decrement from equation 2 for the steady-state condition, substituting V_{qp} for V and P for V_{qp}/E_p, we obtain

$$\frac{P}{(1-P)^{3/2}} = \frac{0.52 f_p \rho}{\beta} \qquad (15)$$

where $\rho = R_2 C/R_1 C$.

This equation was used to compute a table which is similar to a table of logarithms and from which the value of pulse response P can be obtained for any set of values of f_p, ρ, and β. This table is given here as Table I.

Equation 15 can also be solved by the graphical method of Fig. 4 described in the following.

Graphical Method for Determining Pulse Response

Equations 9 and 15 are implicit equations for the quantity P. They do not yield a simple algebraic expression for P in terms of the parameters f_p, ρ, and t_a or β. A graphical solution may be obtained in the following manner.

Either of these equations can be written in the form

$$K(1-P)^n = AP \qquad (16)$$

In this equation P represents the pulse response which is the ratio of quasi-peak reading to peak reading; its value is always less than 1.0. K is a coefficient which is 1.0 for either equation 9 or equation 15. A is another coefficient equal to $2/t_a f_p \rho$ for equation 9 and $1.92\beta/f_p\rho$ for equation 15.

Table II. Respective Values of Pulse Response P from Fig. 4

Rectangular pulse	($n = 1$	$K = 2$)	$P = 0.75$
Rounded pulse	($n = 1.5$	$K = 1$)	$P = 0.51$
Saw-tooth pulse	($n = 2$	$K = 1$)	$P = 0.45$

The exponent n is 2 for equation 9 and 1.5 for equation 15. Equation 5, which gives the value of P for rectangular pulses can be rewritten in the form of equation 16, in which case the exponent n is 1.0. If K is taken as 2.0 the coefficient A becomes $2/t_a f_p \rho$ which is the same as the value of A derived from equation 9.

When both members of equation 16 are plotted against P on the same curve sheet, the point of intersection of the two curves determines P. This is the basis of the chart shown in Fig. 4, which is used in the following manner. The value of A is calculated from the parameters f_p, ρ and t_a or β. A straight line is drawn through A on the right-hand scale and the origin at the lower left, such as the dotted line in Fig. 4. This line represents AP in equation 16. The value of P is then read on the horizontal scale opposite the intersection of the straight line with the applicable curve which represents $K(1-P)^n$. When the value of A exceeds the range of the scale, a straight line is drawn through $1/A$ on the horizontal axis through $A = 1$ on the right-hand scale. Either scale may be used for values of A between 1.0 and 2.0. For example, a straight line which passes through the origin and 2.0 on the right-hand scale intersects the horizontal axis at 0.5 which is the reciprocal of 2.0.

The following example illustrates the use of the chart and also shows the relative response of a quasi-peak voltmeter to pulses of the same peak value, duration, and repetition rate but of different shape. In this example the values of the parameters are $f_p = 30$ pulses per second, $\beta = 6{,}000$-cycle bandwidth, $t_a = 1/6{,}000$ seconds, and $\rho = 600$, which corresponds to time constants of 1-ms charge and 600-ms discharge. The value of A for these parameters is approximately 0.67. The dotted line in Fig. 4 is drawn from point 0 on the left to 0.67 on the right-hand scale. The intersections of this line with the three

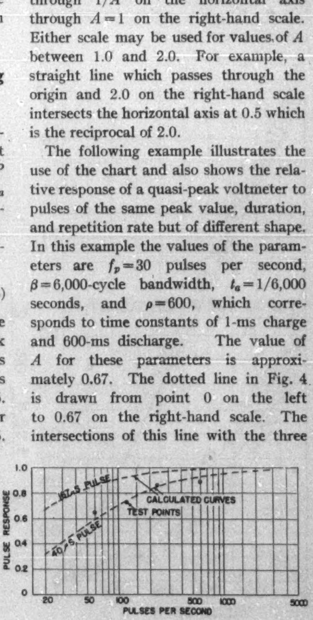


Fig. 6. Response to rectangular pulses, Instrument A, 1-ms charge; 600-ms discharge

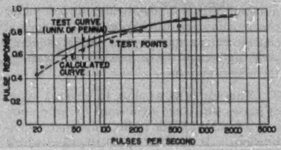


Fig. 7. Response to pulse noise, instrument A, 1-ms charge, 600-ms discharge, 6-kc band

curves on the chart determine the respective values of pulse response P, as shown in Table II.

Since the peak value is the same for all pulses used in this example, the instrument readings for the different pulse shapes will be in the same relation as the values of P. For instance, if the instrument reads 100 on the rectangular pulse, it will read 68 on the rounded pulse and 60 on the saw-tooth pulse.

Comparison of Calculated Values with Test Values of Pulse Response for Rectangular Pulses

Values of pulse response calculated for rectangular pulses and for rounded pulses by the equations given in this paper have been compared with measured values under several different conditions. Curves of response to rectangular pulses plotted against pulse repetition rate are shown in Fig. 6 for 40- and 167-μs pulses. The curves were calculated by equation 5 for an instrument with a discharge-to-charge ratio of 600. Other investigators have analyzed the diode voltmeter.[6]

The test points in Fig. 6 were obtained with a radio noise meter with a quasi-peak measuring circuit in the detector having nominal characteristics of a 1-ms charge and a 600-ms discharge time. This instrument is designated as radio noise meter A in this paper. Rectangular-shaped pulses in the detector output were produced by modulating a radio-frequency signal generator with a pulse generator. The noise meter and the signal generator were tuned to the same frequency. The pulse width was measured with a radio receiver used as a pulse analyzer.[7] The pulse repetition rate was determined by comparison with an oscillator which was calibrated in frequency. Quasi-peak readings were made with the selector switch in the quasi-peak position. Peak voltage was measured by switching to the peak position and adjusting a bias voltage until the audible output of the noise meter was reduced to zero. This is called the slide-back method. The noise meter was then connected to a signal generator and the voltage of this generator was adjusted to give the same instrument scale reading as the bias voltage. The signal generator voltage was then proportional to the peak voltage of the pulse. Since the same calibration applied in both cases, the ratio of quasi-peak reading to peak reading was equal to the pulse response ratio P.

Tests were made only with 40-μs pulses because the maximum pulse width obtainable with the pulse generator used was about 40 μs. The measured width of the pulse used was approximately 41 μs. The agreement between the calculated values and the test values of pulse response for the 40-μs pulses is reasonably good in view of the approximations involved, such as neglecting distortion of pulse shape which may occur in the radio-frequency and intermediate-frequency stages of the noise meter.

When the noise meter used is measuring actual noise pulses or steep wave fronts instead of pulsed carrier, the width of the pulses which appear in the detector circuit is determined by the bandwidth. Since the bandwidth of the instrument was about 6 kc, the pulses would be about 167 μs wide instead of 40. A calculated curve for 167-μs rectangular pulses is shown in Fig. 6 for comparison.

Comparison of Calculated Values with Test Values of Pulse Response for Radio Noise Meters

Pulse response for radio noise under several conditions calculated from Table I was compared with the pulse response measured on four types of radio noise meters in common use. The results are shown in Figs. 7 through 10, on which pulse response is plotted against pulse repetition rate. The noise meters are designated as A, B, C, and D. The test points represent spot checks made by the author on one instrument of each type. Test curves are shown which were obtained in an investigation made at the University of Pennsylvania under contract with the Bureau of Ships, United States Navy. The curves, which were obtained from tests on several instruments of each type, are reproduced by permission of the Bureau from Progress Report No. 22.[8]

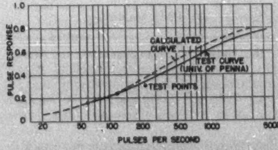


Fig. 8. Response to pulse noise, instrument B, 10-ms charge, 600-ms discharge, 10-kc band

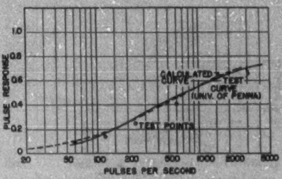


Fig. 9. Response to pulse noise, instrument C, 1-ms charge, 600-ms discharge, 160-kc band

Radio noise meter A, Fig. 7, has characteristics approximating those specified in the proposed ASA standard.[4] The nominal constants are: charge time 1 ms, discharge time 600 ms; bandwidth 6 kc. The radio noise, which was measured in the spot check tests and in the tests at the University of Pennsylvania, was produced by a pulse noise generator built according to a design developed at Purdue University. The amplitude of noise and the repetition rate of the pulses produced by this generator are adjustable over a wide range. The calculations and the test data in Fig. 7 are in reasonably good agreement. In most cases radio noise measurements are not expected to show better than 20-percent accuracy. The calculated values and test values of pulse response in Fig. 7 differ by 10 per cent or less.

Radio noise meter B, Fig. 8, was built according to recommendations of the Joint Coordination Committee in 1940.[8] The nominal constants are: charge time 10 ms, discharge time 600 ms; bandwidth 10 kc. As in the case of noise meter A, the agreement between calculations and test results is reasonably good. The instrument used for the spot checks was not equipped with the slide-back circuit for peak measurements and therefore the peak value had to be determined from the calibrated noise output of the pulse generator and the measured bandwidth of the noise meter.

The effect of the longer charge time of noise meter B may be seen by comparing Fig. 8 with Fig. 7. Noise meter A shows about three times as large a pulse response as noise meter B at 120 pulses per second. The ratio of the readings of the two instruments on the same noise source, however, is less than this because noise meter B has a somewhat larger bandwidth and the peak reading is directly proportional to bandwidth. Allowing for the bandwidths of 6 and 10 kc

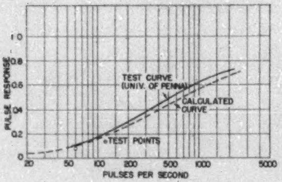


Fig. 10. Response to pulse noise, instrument D, 1-ms charge; 600-ms discharge; 160-kc band

respectively, the ratio of instrument readings becomes about 2 to 1.

Radio noise meters C and D have characteristics approximating those specified in another proposed ASA standard,[6] which applies to frequencies above 20 megacycles. The nominal constants are: charge time 1 ms; discharge time 600 ms; bandwidth 160 kc. Fig. 9 applies to noise meter C and Fig. 10 to noise meter D. As in the case of the other two instruments, the agreement between calculations and test data is reasonably good. The instrument D used for the spot checks was not equipped for making peak measurements by the slide-back method. Peak values were obtained from the calibration of the noise generator and the measured bandwidth of the instrument as in the case of noise meter B. This procedure did not allow for attenuation in the matching pad, which would have been detected by the slide-back measurement. This is probably the reason for the relatively low values of pulse response shown by the test points in Fig. 10.

Calibration of the Quasi-Peak Voltmeter

In the basic circuit of Fig. 1 a measurement of the average voltage on the capacitor C in the steady-state condition is required. It is assumed that the presence of the voltmeter does not affect the resistance R_2. In most radio noise meters the indicating instrument is a d-c voltmeter which is slightly underdamped and has a response time of about 300 ms. The response time is the required time to come to apparent rest after an abrupt change in the applied voltage. The voltmeter which might include the resistance R_2 in its circuit could be calibrated with direct voltage by itself or through the input terminals shown in Fig. 1. In the latter case it is theoretically necessary to allow for the voltage drop in resistance R_1. This is accomplished by multiplying the input voltage by $\rho/\rho+1$ where ρ is the ratio of discharge time to charge time. In most cases the value of ρ is at least 50, in which case the voltage drop in R_1 is less than 2 per cent.

In the case of the radio noise meter the effect of amplification in the radio-frequencies and intermediate-frequency stages must be included in the calibration. This is accomplished by impressing a sine wave voltage of the proper radio frequency on the input terminals of the noise meter. The detector output as applied to the voltmeter is substantially a direct voltage equal to the applied peak voltage multiplied by an amplification factor. This calibration is usually made in terms of rms voltage because radio signals are usually measured in terms of the rms voltage of the carrier. To obtain the actual voltage impressed on the circuit of Fig. 1 the signal generator rms voltage is multiplied by 1.41 times the amplification factor. It should be understood that the rms voltage used for calibration has no direct relation to the rms value of pulse voltage at the detector which is discussed in the Appendix.

Radio noise meters are usually calibrated in rms microvolts for either quasi-peak or peak readings. The factor 1.41 cancels out when the pulse response, which is the radio of quasi-peak to peak, is calibrated.

Conclusions

1. It is practical to calculate the relative response to pulse type noise of different noise meters which measure quasi-peak value.
2. Approximate calculations based on the nominal values of bandwidth and detector charge time and discharge time for several representative noise meters show reasonably good agreement with test data taken on those instruments.
3. This analysis applies to several types of pulses which are significant to radio noise and are suitable for radio noise meter calibrations. It does not apply to pulses which are unequal or random with respect to magnitude, recurrence rate, or both.
4. This investigation contributes the first of the two steps necessary to the interpretation of radio noise meter readings in terms of radio noise nuisance value. The first step correlates characteristic noise wave shapes with noise meter readings and this step is now quite clearly taken as disclosed in this paper. The second step correlates noise meter readings with annoyance to a listener. This second step should be given more attention at this time.

Table III. Comparison of Quasi-Peak, Average, and RMS Values

	Peak	Quasi-Peak*	Average	RMS
Square wave, half-cycles	1.00		1.00	1.00
Sine wave, half-cycles	1.00		0.637	0.707
Envelope of 50 per cent modulated Radio frequency wave, 120 cycles modulation frequency	1.50	1.47	1.00	1.06
167-μs rectangular pulse 120 pulses per second	1.00	0.92	0.02	0.14
40-μs rectangular pulse 120 pulses per second	1.00	0.74	0.0048	0.07

*Circuit of Fig. 1 with 1 ms for charging and 600 ms for discharging.

Appendix. Quasi-Peak, Average, and Rms Values for Several Types of Waves

Relation Between Quasi-Peak Voltage and Average of Pulsating Voltage

Most radio noise meters provide for switching the detector output to a circuit for measuring "average" value. This circuit is normally used for measuring radio signals, in which case the effect of modulation is eliminated and the radio-frequency carrier is measured. It has been suggested that average readings of pulse noise might be useful to supplement the quasi-peak readings. Therefore, it may be of interest to calculate the relation between average voltage and quasi-peak voltage as illustrated here for the case of rectangular pulses such as those of Fig. 2. The average voltage is equal to the area of the pulse divided by the time of 1 pulse cycle. The time of 1 pulse cycle is $1/f_p$. Then the average voltage is given by the equation

$$V_{av} = E_p t_a f_p \tag{17}$$

Average values for different waves are compared in Table III. The ratio of average voltage to quasi-peak voltage is obtained by combining equation 17 with equation 4 which gives the relation

$$\frac{V_{av}}{V_{qp}} = t_a f_p + \frac{1}{\rho} \tag{18}$$

When the pulse width is short compared with the time of 1 pulse cycle the average reading is very small. For example, when measurements are made with an instrument having a 1-ms charge time and a 600-ms discharge time for the quasi-peak measurement, and a 40-μs pulse repeated 120 times per second is measured, the ratio of average to quasi-peak is 0.0065.

Relation Between Quasi-Peak Voltage and RMS of Pulsating Voltage

Although presently available radio noise meters do not measure the rms voltage in

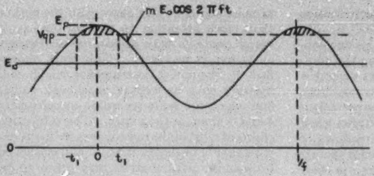


Fig. 11 (left). Envelope of detector output with sine wave modulation

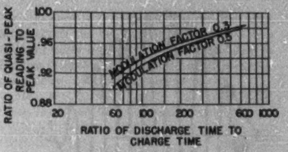


Fig. 12. Response to quasi-peak voltmeter to radio frequency with sine wave modulation. The curves apply to any modulation frequency in the range of 20 to 1,000 cycles

the detector output, it may be of interest to calculate the ratio of rms voltage to quasi-peak voltage. The measurement of rms detector voltage would require a special type of voltmeter since the dynamometer type is not suitable for this application. It does not appear at this time that such an addition would serve a useful purpose.

The rms value of the rectangular pulse is obtained by plotting the curve of Fig. 2 with the voltage squared, dividing the area by the time of 1 pulse cycle, and taking the square root of the quotient. Expressed as an equation, this becomes

$$V_{rms}=E_p\sqrt{t_a f_p} \qquad (19)$$

where $1/f_p$ is the time of 1 pulse cycle.

Rms values for different waves are compared in Table III. The ratio of rms value to quasi-peak value from equations 4 and 20 is

$$\frac{V_{rms}}{V_{qp}}=\sqrt{t_a f_p}+\frac{1}{\rho\sqrt{t_a f_p}} \qquad (20)$$

In the case of the 40-μs pulse with a 120-cycle repetition rate used to illustrate the average value, the ratio of rms voltage to quasi-peak voltage is 0.093 when $\rho=600$ for the quasi-peak measurement.

Quasi-Peak Value of Sine Wave Modulated Radio-Frequency Voltage

The quasi-peak circuit of the radio noise meter is not ordinarily used to measure modulated signals. However, it may be useful to know the effect of modulation at audio-frequencies such as 400 cycles. An analysis of the effect of such modulation on quasi-peak readings is given here.

The envelope of the output of the rectifier *A* in Fig. 1 when a modulated high-frequency voltage is applied to the input terminals is shown in Fig. 11. When the steady-state condition is reached, the rectifier output voltage exceeds the capacitor voltage during the interval $-t_1$ to t_1 where the zero point is at the peak of the modulated wave. During this interval (see shaded area in Fig. 11) the capacitor receives a charge to make up for the loss during the discharge period, which is approximately equal to $1/f$ where f is the modulation frequency.

Proceeding as in the analysis of pulses and equating the voltage increment to the voltage decrement we obtain a relation between quasi-peak readings and peak readings. If P_m represents the response to modulation which corresponds to the factor P for pulses, we obtain the equation

$$\frac{P_m}{(1-P_m)^{3/2}}=0.3\rho\sqrt{\frac{1+m}{m}} \qquad (21)$$

In this equation ρ is the ratio of discharge time to charge time as before and m is the modulation factor which ranges from zero to 1.0. Equation 21 becomes indeterminate when m is zero but gives definite results for the usual modulation factors of 0.3 and 0.5. For smaller modulation factors conditions approach those of an unmodulated carrier wave which is discussed in the section on calibrations.

The similarity may be noted between equations 21 and 15 for rounded pulses. The graphical method of Fig. 4 may be used to obtain values of P_m. The curve for $n=1.5$ is used and the factor A is the reciprocal of the right-hand member of equation 21. For this case it will be necessary to enlarge the chart or at least the lower right-hand portion.

The curves of Fig. 12 cover the full range of conditions likely to be encountered in most situations. It may be noted that the quasi-peak reading is within 10 per cent of the peak value and that there is little difference between the curves for 0.3 and 0.5 modulation (30 to 50 per cent). Modulation frequency within the range of about 20 to 1,000 cycles is not a factor. In the case of the radio noise meters considered in this investigation, the value of ρ is 600 for noise meters *A*, *C*, and *D*, and 60 for noise meter *B*. Measured values of P_m for noise meters *A* and *B* using 0.3 modulation agree with the calculated values within the accuracy of the measurements.

References

1. Measurements Pertaining to the Coordination of Radio Reception With Power Apparatus and Systems, C. M. Foust, C. W. Frick. *AIEE Transactions (Electrical Engineering)*, vol. 62, June 1943, pp. 284–91.

2. A Study of Wave Shapes for Radio Noise Meter Calibrations, C. W. Frick. *AIEE Transactions*, vol. 64, 1945, pp. 890–901.

3. Methods of Measuring Radio Noise, Joint Coordination Committee on Radio Reception of Edison Electric Institute—National Electrical Manufacturers Association—Radio Manufacturers Association Report. *EEI Publication No. G9, NEMA Publication No. 107, RMA Engineering Bulletin No. 32*, New York, N. Y., 1940.

4. Proposed American Standard for a Radio Noise Meter, 0.015 to 25 Megacycles/Second. *ASA Publication No. C63.2, NEMA Publication No. 102-1950, RTMA Engineering Bulletin No. 32-A*, New York, N. Y., 1950.

5. Proposed American Standard for a Radio Noise and Field Intensity Meter, 20 to 1,000 Megacycles/Second. *ASA Publication No. C63.3, NEMA Publication No. 131-1952, RTMA Engineering Bulletin No. 41*, New York, N. Y., 1952.

6. Report of the Meeting of the Group of Experts. *Report Number (R.I.)5*, International Electrotechnical Commission, International Special Committee on Radio Interference, London, England, May 19–22, 1936.

7. Pulse Response of Diode Voltmeters, A. Easton. *Electronics*, New York, N. Y., vol. 19, Jan. 1946, pp. 146–49.

8. The Communications Receiver as a Wave Analyzer, Arnold Peterson. *General Radio Experimenter*, Cambridge, Mass., vol. 18, Jan. 1944, pp. 1-4.

9. Investigation of the Measurement of Noise. *Progress Report No. 22*, University of Pennsylvania, Philadelphia, Pa.

◆

Discussion

Alfred Eckersley (Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pa.): The state of the art of radio-interference measurement has been advancing considerably in the last few years, but published information on the subject is still very meager. Any contribution to the literature is therefore very welcome.

At the Moore School, we have made an an analysis of the quasi-peak detector circuit response to pulses, as shaped by the band-pass characteristic of a radio-noise meter. This was done both by a graphical method and an analytical one. The latter was based partly on the work of Burgess.[1] The solutions we obtained have a form very similar to those given in the paper. It is worth while pointing out, however, that certain improvements can be made on the method of solution used in this paper which will improve the accuracy of the final result.

First, the author has taken the envelope function of the pulse as the voltage which causes the diode to conduct. Actually, the sinusoid within the envelope should be considered the driving function; this would cause a considerable difference in the amount of charge acquired by the capacitor during the conducting period.

Second, the envelope function (Fig. 5) is not the correct one. The response of casacaded tuned circuits to an impulse has an envelope that is similar to a normal distribution curve, rather than the sin x/x curve, which applies to an idealized rectangular filter. If these factors are taken into account, Fig. 5 would be replaced by Fig. 13. This wave form has been observed in noise meters at the detector input. The effective width of this pulse envelope i.e., the width of a rectangle with the same

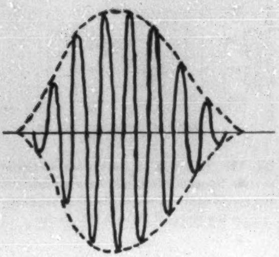


Fig. 13. Fig. 5 revised

peak value) is approximately the reciprocal 6-decibel bandwidth of the meter. It should be noted that if the actual envelope is replaced for convenience in analysis by some other shape there is necessarily some error involved. Replacement by the equal area pulse does not make the error zero, or necessarily small.

The inverse bandwidth of the meter is the minimum plus envelope duration that can be obtained at the detector input. Excitation with long rectangular carrier pulses will produce the rectangular envelope considered by the author; the rise time is, however, still dependent on the bandwidth. Thus, in Fig. 6, the 167-μs curve is the more correct one to apply. The reason for disagreement with the measured values is thought to be the approximations employed in this paper.

A clarification of the terminology used in connection with quasi-peak detector time constants, which are mentioned in the beginning of the paper, may be in order. The ASA Specification *C63.2* (ref. 4 of the paper) requires a 1-ms charge time and a 600-ms discharge time for the quasi-peak detector in a noise meter. In terms of the quantities on Fig. 1, the charge time may be shown to be approximately 3.6 R_1C; it is defined as the time taken for the rectified voltage to reach 63 per cent of its final value after the application of a step sinusoid to the detector. This contrasts with the charge time constant R_1C, the factor of 3.6 being obtained by consideration of the fact that the detector only charges during positive half-cycles of the sinusoid, according to a law which is only approximately exponential. The discharge time constant is R_2C. Thus, for a 1-ms–600-ms circuit, R_2/R_1 is approximately 2,000.

In conclusion, I would agree strongly with the author that an important problem remaining in this field is the scientific investigation of the relation between noise meter readings and relative interference effect. The difficulties in the attack on this problem are large. They include the diversity of communications systems (and their interference susceptibilities) and the question of quantitative definition of interference with the intelligence being received.

REFERENCE

1. THE RESPONSE OF A LINEAR DIODE VOLTMETER TO SINGLE AND RECURRENT IMPULSES OF VARIOUS SHAPES, R. F. Burgess. *Proceedings*, Institution of Electrical Engineers, London, England, vol. 95, pt. III, March 1948, pp. 106–10.

James M. McCutchen (Rural Electrification Administration, Washington, D.C.): The author is to be congratulated on his development of the method of correlating the readings of the different noise meters on pulse-type noise. The variations in wave shape have led to errors in meter interpretation; this has caused considerable confusion when attempts are made to compare readings of different instruments. Two major points are yet to be covered, however, in that many types of pulse interference which originate particularly in power-line apparatus are unequal or random with respect to magnitude and repetition rate. A rod-to-plate type of discharge particularly shows this random magnitude effect as the ionization path transfers around the outside end of the rod because of magnetic and thermal effects in the arc. Further effort might be made in this direction to solve some of the problems in comparing measurements on these types of noise sources.

Recordings which we have made on this type of noise on variable area movie sound track show these pulse effects which vary both in rate and amplitude. Of course, the galvanometer mirror used in the recording apparatus has some effect in overshoot of these pulses, but not to the extent shown on the recordings. This type of random variation makes reading comparisons difficult. These variations will have considerable bearing too on the evaluation of the nuisance value of noise to the listener which needs further investigation; the author has pointed out the need of further work in this field. I believe that this part of the investigation might include the effects of the poor transient response and overshoot of the audio section and loudspeaker of the home receiver where the damping of these speakers is considerably less than ideal. These effects tend to exaggerate the pulse-type interference over that of sine-wave interference and may lead to assigning a higher "nuisance multiplier" to the pulse type of interference.

W. E. Pakala and **D. C. Fahrnkopf** (Westinghouse Electric Corporation, East Pittsburgh, Pa.): This paper will be very useful to those making radio noise measurements. It will also be eductional in that the quasi-peak function of radio noise meters which is considered and analyzed here has not been sufficiently publicized.

It is one purpose or objective of radio noise studies to be able to correlate radio noise instruments and measure noise so that results can be used to determine the interference effects to communication. With instrument *A* of the paper it is possible to measure both the quasi-peak and peak values and to obtain a measured pulse response for any particular type of radio noise. Unfortunately, no analytical means is available as yet for obtaining the net effect on a particular form of communication.

Only by comprehensive listening and viewing tests can the effect or nuisance value of any one type of noise be determined. Listening tests were made several years ago in the United States and Canada with meters having various time constants, and viewing tests have been made in Canada within the last 3 years using several types of noise sources. We agree with the author that more attention to the correlation of meter readings with annoyance to listener or viewer is necessary. This work should include the latest instruments which will measure peak as well as quasi-peak values and which are built according to references 4 and 5 given in the paper. The subjective effects of most of the radio noise producing devices should be determined. Comprehensive tests and careful analysis may prove to be quite illuminating and narrow down the problem more than one would expect.

The ratio of instrument reading on the two noise meters *A* and *B* of about two to one has been our experience when measuring interference from a 60-cycle source where noise pulses on 1/2 cycle give 120 pulses per second. The average of many readings has given a ratio slightly higher, about 2.1, which checks very well.

At very low or intermittent pulse rates, it is almost impossible to obtain a consistent usable reading because of meter fluctuations on quasi-peak. Under these conditions it is preferable to read the peak value in order to better evaluate the radio noise source or the degree of interference.

H. L. Rorden (Bonneville Power Administration, Portland, Oreg.): The quasi-peak voltmeter described is of particular interest and is a very welcome contribution to investigators who have spent much time and effort attempting to better understand the causes and effects of radio noise, as well as other phenomena associated with corona. The difficulties of obtaining realistic test data in the study of radio noise have been further increased by lack of an adequate instrument to be assured that results are comparable to conditions in service. Not only have instruments been limited in their ability to read crest voltages at the higher frequencies but because of their sensitivity they have been subject to a wide tolerance, resulting in data that are difficult to pinpoint. If indicated data differ from the true audible noise produced in radios, a false sense of security may result.

This paper should be of considerable value to the power transmission equipment engineer in assisting his analyses of readings taken with one of the radio noise meters. It should also receive serious consideration in committees reviewing standards for the measurement of radio noise.

C. W. Frick: I wish to thank the discussers for their contributions. It is gratifying to note that they recognize the usefulness of the data presented and the need for further work on the correlation of radio noise measurements with the effects on communication.

Mr. Eckersley questions the use of the envelope function to represent the detector output of the radio noise meter. If we disregard the fine structure, the envelope shows the pulse shape. One purpose of the paper was to show the effect of pulse shape on the quasi-peak reading. If pulses of the shapes shown by Figs. 2, 3, and 5 are applied to the circuit of Fig. 1, the ratio of quasi-peak to peak is as calculated. The application of this analysis to the radio noise meter is, of course, subject to other considerations.

There is some similarity between radio noise and telephone noise that may be caused by power apparatus. In the early days it was considered almost hopeless to calculate telephone noise. However, by taking the most significant factors, comparing calculations with test data, and including other factors as necessary, "short-cut" methods have been developed which are useful at least for estimating purposes. Radio noise seemed even more difficult to treat by calculation. This was not because there were more factors but rather because the factors involved were not so well understood. Again, by taking the most significant factors the radio noise produced by a square wave generator or a pulse generator has been calculated approximately.

At first the detector output of the radio noise meter was treated as a rectangular pulse. There was apparent agreement with tests with the earlier radio noise meters which had the slower charge time. When the newer instruments which have the faster charge time were tried, there was disagreement between calculations and tests, and the calculation was improved by using the pulse shape shown in Fig. 5. Mr. Eckersley's discussion indicates that an analysis has been made which takes into account other factors. These results should further improve the method and it is hoped that they will be published soon.

Mr. McCutchen points out that radio noise produced by power apparatus may be subject to irregularity as to magnitude and repetition rate. I believe that even so the concept of pulse noise helps us to understand the causes of radio noise. However, it should be realized that in the present state of the art the evaluation of radio noise is not an exact measurement.

An Operational-Type Magnetic Amplifier for Air-borne Servo-Control Systems

R. M. HUBBARD
NONMEMBER AIEE

THIS paper describes a magnetic amplifier developed to meet the need for a versatile, reliable, easily reproducible operational component for a low-level servo-control system. The subject amplifier also exhibits unusual stability, linearity, and freedom from drift, which are requirements of prime importance in this particular application.

The most useful feature of this circuit is that it satisfies, within the limitation of its passband, the requirements for an operational element in feedback control systems, i.e., it can be used to perform various functions or operations without imposing its own characteristics on the system. It has made possible the design of a single basic amplifier with fixed standard components and it is used to perform such varied functions as mixing signals, changing impedance, adjusting gains, and driving hydraulic control valves. This basic "building block" amplifier lends itself admirably to quantity production techniques. Sufficiently fast response and light weight have been achieved through the use of a 900-cycle power source, which in this particular application is specified by the requirements of associated subsystems. Output powers of the order of 100 milliwatts with power gains of 200 to 300 are typical.

Basic Theory of Operation

The so-called operational magnetic amplifier as developed for air-borne servo-control systems is shown in its most basic form in Fig. 1, and in equivalent block diagram form in Fig. 2(A). This circuit utilizes a familiar basic principle first applied to magnetic amplifiers by Geyger in his self-balancing potentiometer type of circuit.[1] The addition of compound feedback and a novel stabilized self-bias circuit N_b to a basic self-saturating amplifier produces an amplifier having operating characteristics not normally obtainable with magnetic circuits.

The use of positive feedback gives this amplifier an effectively infinite internal gain which is stabilized by an adjustable external negative feedback. As a result of this compound feedback arrangement the control winding N_c has only a transient control over the circuit. Consider, for example, the application of a control current I_c which causes a control coil current to flow, producing a differential output voltage. This output produces a negative-feedback current which is introduced directly into the control winding through R_s to buck out the control current. In the final equilibrium condition the control coil current is essentially zero and a new input and output level is established, with the d-c control ampere-turns necessary to maintain this level supplied by the positive feedback winding N_f. It is apparent that if the control coil current is zero, the input voltage is zero and the output voltage E_0 is given by the relationship

$$E_0 = I_c R_s \tag{1}$$

It can therefore be seen that the gain of the circuit is a linear function of R_s and the input impedance is effectively zero. The fact that the output voltage is regulated to $I_c R_s$ makes it virtually independent of load impedance. A simple change in the negative feedback configuration, Fig. 2(B), permits the output current I_0 to be regulated such that

$$I_0 = I_c(R_s + R_p)/R_p \tag{2}$$

In this case the output current is virtually independent of load resistance.

It is the combination of extremely low input impedance and a linear gain function that has prompted the classification of this device as an operational amplifier. It can be seen from Fig. 2(A), e.g., if the input impedance is effectively zero

$$E_0 = R_s(I_1 + I_2 + I_3 + \ldots) \quad \text{or,} \tag{3}$$

$$E_0 = R_s(E_1/R_1 + E_2/R_2 + E_3/R_3 + \ldots) \tag{4}$$

The fact that the basic amplifier without the negative feedback path provided by R_s is truly an infinite-gain device is interesting as it represents the "ideal" 100-per-cent feedback magnetic amplifier. This suggests that the conventional self-saturating circuit, wherein the gain is less than infinite owing to an effective negative feedback which is dependent on the magnetic characteristics of the reactor cores, leakage reactance, rectifier leakage, etc., may reasonably be considered a "nominal" 100-per-cent feedback amplifier. In effect, the positive feedback around the basic amplifier unit in the operational circuit is adjusted to compensate for these variable factors which

---

Paper 54-247, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 23, 1954; made available for printing April 26, 1954.

R. M. HUBBARD is with the Boeing Airplane Company, Seattle, Wash.

The author wishes to acknowledge with thanks the guidance and encouragement of Dr. Kenneth D. Johansen, whose assistance made this work possible.

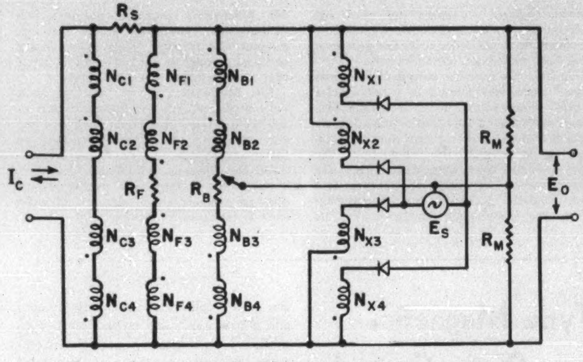


Fig. 1. Basic operational amplifier circuit diagram, voltage output type

would otherwise decrease the maximum available gain. The subsequent addition of negative feedback to the basic infinite-gain amplifier permits predictable operation at all gains from infinity to zero.

Nomenclature

E_o = average differential output voltage
E_s = effective power supply voltage
I_c = average control (signal) current
I_o = average differential output current
N_b = number of turns in bias winding
N_c = number of turns in control winding
N_f = number of turns in positive feedback winding
N_x = number of turns in load winding
R_b = bias adjusting resistor
R_f = positive feedback adjusting resistor
R_g = signal generator resistance
R_m = mixing resistor
R_p = negative feedback current-dividing resistor
R_s = negative feedback adjusting resistor

General Operating Characteristics

In a practical circuit the idealized operating characteristics can only be approached. As a result of the inherent nonlinearities in the reactor characteristics, the positive feedback winding N_f cannot supply exactly the ampere-turns required to attain an infinite internal gain over the entire working range of the amplifier. Since it is necessary for the control winding to provide the residue of ampere-turns, it is therefore impossible for the control coil current to be exactly zero at all output levels. The error in applying equations 1 and 2 can be predicted on the basis of the ratio of control coil current to control current I_c. The allowable error imposes a practical limitation on the maximum usable gain. In fact, as is the case with any simple self-saturating circuit, the stability and linearity of the operational circuit are directly dependent on the amount of negative feedback.

Although the gain of this circuit is independent of the number of control turns, it is desirable in the interest of optimum linearity to use as many turns as the required response time will permit. In this way, the inevitable residue of compensating ampere-turns which must be supplied by the control circuit will require a minimum of current from that source, resulting in a minimum error in the application of equations 1 and 2. The use of the greatest possible number of control turns has the important added advantage of improving the stability by providing for a maximum of negative feedback ampere-turns.

This amplifier is extremely stable and quite insensitive to sizable changes in supply voltage and frequency, and to large temperature variants. This stability is achieved by two methods. First, the amplifiers are usually operated with sufficient negative feedback to fix the gain at a value less than that of the basic self-saturating circuit. The fact that the negative feedback is derived from a differential output makes its use especially advantageous in the reduction of zero drift. Secondly, the self-bias is derived from the negative feedback stabilized voltage across each mix resistor R_m and therefore its dependence on supply regulation and changes in rectifier leakage is very small. Finally, there is a negative feedback inherent in the biasing action. Any tendency toward a change in the output of either half of the circuit produces a bias opposing such a change. This negative feedback operating on the individual halves of the circuit also minimizes to a great extent the dissymmetry normally introduced by mismatch between pairs of cores. The nonlinearities usually introduced by mismatch between the individual cores in either half of the circuit are greatly reduced by the addition of a shunt resistor across each pair of bias windings to provide a low-impedance coupling path.

These two effects, the inherent negative feedback in the bias circuit and the action of the coupling circuit, are illustrated in Fig. 3. The transfer characteristic curve A was obtained with poorly matched cores in a simple self-saturating

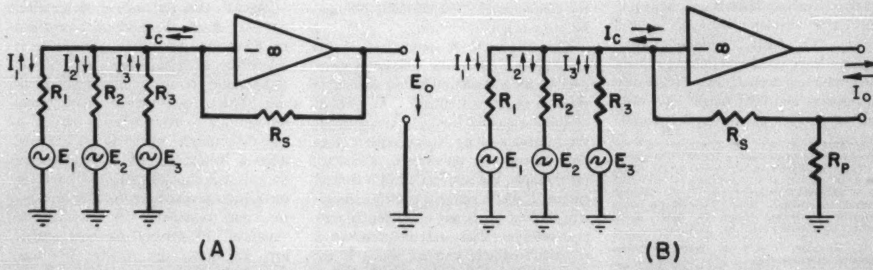


Fig. 2. Operational amplifier block diagram

A—Voltage output type

B—Current output type

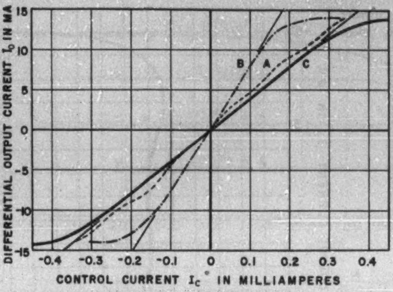
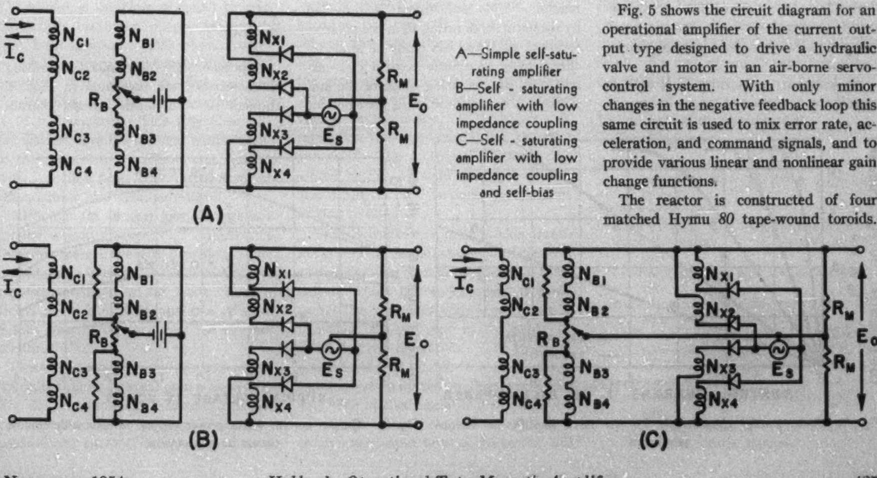


Fig. 3 (above). Transfer characteristics for a self-saturating amplifier having mismatched cores

A—Fixed bias
B—Fixed bias and low-impedance coupling
C—Self-bias and low-impedance coupling

circuit with fixed bias; see Fig. 4(A). The addition of a low-impedance coupling path, Fig. 4(B), yielded the improvement in linearity illustrated by curve *B*. A self-bias circuit was added, Fig. 4(C), which corrected the dissymmetry evident in curve *B* and resulted in the linear, symmetrical transfer characteristic curve *C*.

As a result of all these compensating features the output of this type of amplifier has, for most applications, an almost negligible dependence on core matching, rectifier leakage, supply voltage and frequency, and temperature (within the limitations of the rectifiers). It is further possible with this circuit to approach closely the transfer characteristic of an ideal amplifier, i.e., good linearity with a sharp saturation and an essentially constant value thereafter.



(A)

(B)

(C)

Fig. 4 (below). Circuit diagrams of amplifiers used in illustrating compensation for core mismatch

A—Simple self-saturating amplifier
B—Self-saturating amplifier with low impedance coupling
C—Self-saturating amplifier with low impedance coupling and self-bias

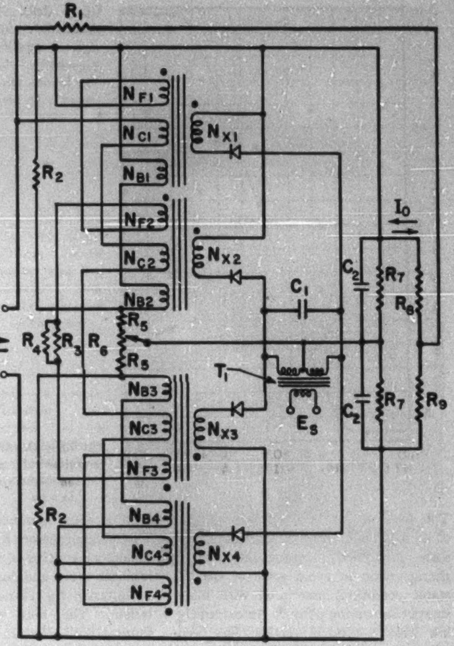


Fig. 5. Circuit diagram of a magnetic operational amplifier to drive an electrohydraulic control valve

An Application

Fig. 5 shows the circuit diagram for an operational amplifier of the current output type designed to drive a hydraulic valve and motor in an air-borne servo-control system. With only minor changes in the negative feedback loop this same circuit is used to mix error rate, acceleration, and command signals, and to provide various linear and nonlinear gain change functions.

The reactor is constructed of four matched Hymu *80* tape-wound toroids.

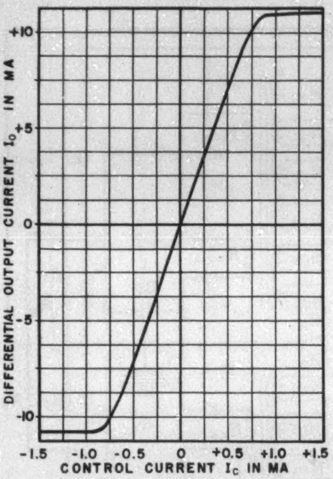


Fig. 6 (left). Static transfer characteristic for the valve amplifier

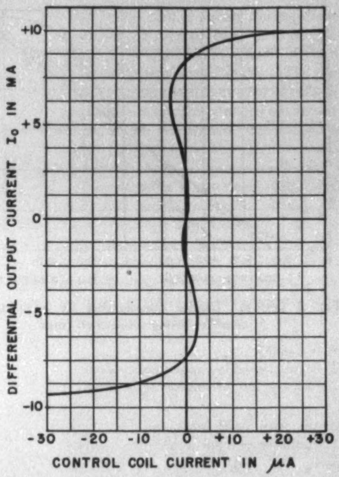


Fig. 8 (right). Output current versus control coil current for the valve amplifier

This material has been selected because of its low magnetizing ampere-turns and more rounded magnetization-curve characteristic to avoid some of the bistable conditions associated with more sharply saturating alloys. An extremely low leakage type of rectifier has been selected to minimize the effect of changes in leakage with temperature. While the new type of self-bias makes the circuit insensitive to sizable changes in rectifier leakage, the zero drift requirements in this particular application are so stringent that every source of possible drift has to be minimized. In fact, it has been found desirable to age the rectifiers at an elevated temperature while passing rated current to stabilize their leakage characteristics.

An isolation transformer T_1 is used to permit the operation of stages in cascade from a single grounded power supply. The mixing resistors R_7 are shunted with capacitors to give a more symmetrical response to rise and decay transients, and to linearize the static transfer characteristic. The bias windings N_b are shunted with resistors R_2 to provide the coupling necessary for linearity and a resultant good infinite-gain characteristic. Since the positive feedback required to attain an infinite internal gain is dependent on the slope of the basic self-saturating transfer characteristic, some manufacturing adjustment is necessary to take into account normal differences between reactors. This adjustment is provided by means of R_4 shunting R_3 in the positive feedback path and can easily be made by checking for minimum control coil current over the operating range of the amplifier. A potentiometer R_6 provides for zero balance while R_5 fixes the bias level at the proper value. The current gain of the circuit, as determined from equation 2, is fixed by R_1 and R_9 such that

$$I_o/I_c=(R_1+R_9)/R_9 \tag{5}$$

The electrohydraulic control valve is represented by R_8 in the circuit diagram. Since the copper valve-windings are immersed in hydraulic fluid, their resistance is a function of the fluid temperature. The fact that this load resistance is a variable makes the current output-type operational amplifier ideal in this application.

Results

Experimentally determined operating characteristics are presented in Figs. 6 through 12. The static transfer charac-

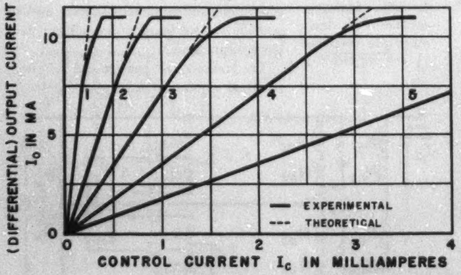


Fig. 7. Static transfer characteristic for the valve amplifier at various gain levels

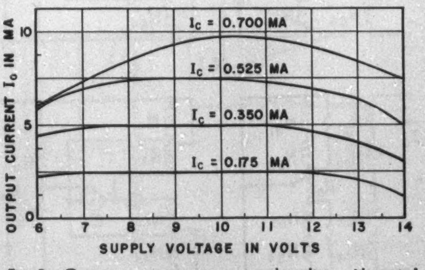
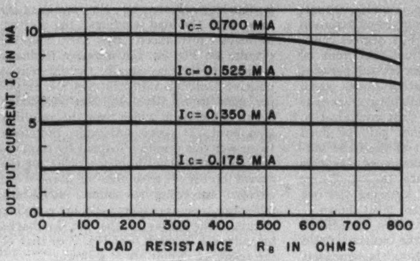


Fig. 9. Output current versus power supply voltage with control current as a parameter



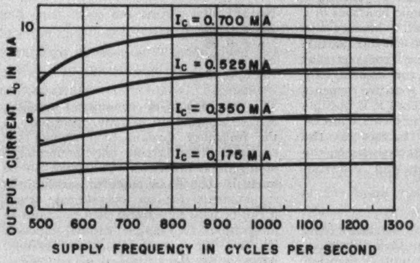


Fig. 10 (left). Output current versus power supply frequency with control current as a parameter

Fig. 11 (left). Output current versus load resistance with control current as a parameter

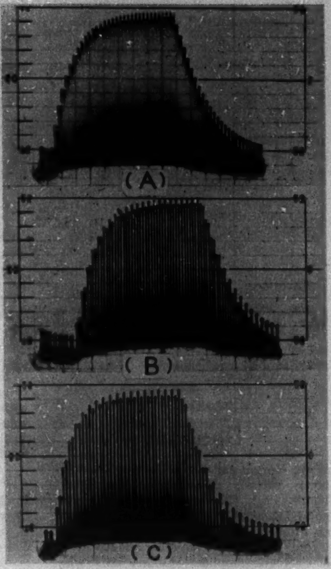


Fig. 12 (right). Transient response to a square voltage pulse with signal generator resistance of

A—500 ohms
B—1,000 ohms
C—2,000 ohms

teristic for this circuit with a load resistance R_s of 500 ohms is shown in Fig. 6. In this application it is necessary to have a sharp saturation, with the output remaining very nearly constant at the saturation level in the presence of large oversignals. Actually the output drops less than 10 per cent with an oversignal of 10 times that required for saturation. It should be noted that although the gain of the circuit is very nearly independent of load resistance the saturation level is not. The circuit is therefore designed to give the proper maximum output level with that load resistance, 500 ohms, corresponding to normal hydraulic fluid temperature. The design current gain of this circuit is 14.3 and the saturation level is 11 milliamperes. This typical amplifier has less than 0.4-per-cent nonlinearity in the range 0 to 7 milliamperes' output with less than 6-per-cent nonlinearity at the nominal 10-milliampere level.

Although the original circuit does not require a gain change, Fig. 7 illustrates that this feature is easily provided. The dotted lines indicate theoretical characteristics predicted on the basis of equation 5 in which R_1 is the variable. Curve 5 falls exactly on the predicted characteristic.

The fact that this amplifier has an essentially infinite internal gain is shown by Fig. 8. The input impedance of this circuit is dependent on the magnitude of the control coil current, and is less than 1 ohm in the range 0 to 9 milliamperes' output. In other applications the working range of the amplifier is limited to a smaller portion of the total dynamic range and an even more nearly ideal infinite-gain characteristic is attained.

Figs. 9 and 10 show the effect of supply regulation on the amplifier, wherein normal supply voltage and frequency are 10 volts and 900 cycles per second. Perhaps the most important point to be considered here is the effect of supply regulation on the zero balance. In this circuit the zero error is less than 0.2-per-cent with 10-per-cent regulation of supply voltage or frequency. With 20-per-cent regulation the zero error is less than 0.4 per cent. The experimental data presented in Fig. 11 show that the current gain of the circuit is virtually independent of load resistance.

The allowable zero drift in this amplifier is very small and, as previously stated, every effort has been made to design a circuit in which this requirement is satisfied. The zero drift with temperature is less than 0.1 per cent in the range 20 to 60 degrees' centigrade. The transient response of this circuit to a step-voltage input is shown by the oscillograms in Fig. 12. The applied signal is a square pulse of voltage whose peak value is sufficient to give an output level equal to one-half of the maximum. It has been determined that in this range the transient variation in peak output current is a good measure of the variation in average output current. In Fig. 12(A), the signal generator resistance R_g is 500 ohms, and the 63-per-cent rise and decay times are respectively 6 cycles and 9 cycles of the 900-cycle carrier. When R_g is increased to 1,000 ohms, Fig. 12(B), the rise and decay times become 4 and 6 cycles respectively. A further increase to 2,000 ohms, Fig. 12(C), results in a rise time of 3 cycles and a decay time of 4 cycles.

Valve Amplifier Component Specifications

Reactor cores: Hymu 80 0.001-inch tape-wound toroids, 1/2 by 3/4 by 1/8 inches
Reactor windings:
  Load winding, N_x—600 turns, no. 36 wire
  Control winding, N_c—400 turns, no. 36 wire
  Feedback winding, N_f—100 turns, no. 36 wire
  Bias winding, N_b—100 turns, no. 36 wire
Transformer core: Hipernik V 0.002-inch tape-wound toroid, 5/8 by 3/4 by 3/16 inches
Transformer windings:
  Low-voltage winding—400 turns, no. 34 wire
  High-voltage winding—1,400 turns, no. 34 wire, center tapped

Rectifiers: One-cell selenium, 30 milliamperes rated, Westinghouse *12G11-KE1C*

R_1: 1,330 ohms
R_2: 180 ohms
R_3: 4,700 ohms
R_4: 2,700 ohms, chosen experimentally
R_5: 3,000 ohms
R_6: 500-ohm potentiometer
R_7: 220 ohms
R_8: hydraulic valve coil resistance, nominally 500 ohms
R_9: 100 ohms
C_1: 0.5 microfarad
C_2: 2 microfarads
E_2: 10-volt 900-cycle power supply

Reference

1. MAGNETIC AMPLIFIERS OF THE SELF-BALANCING POTENTIOMETER TYPE, W. A. Geyger. *AIEE Transactions*, vol. 71, pt. I, 1952 (Jan. 1953 section), pp. 383–95.

◆

Discussion

W. F. Horton and **R. O. Decker** (Westinghouse Electric Corporation, East Pittsburgh, Pa.): We should like to express appreciation to the author for presenting an interesting and useful paper. In part it confirms the results of other investigations of magnetic amplifier stability over wide environmental ranges. In general we have found that the full-wave push-pull self-saturating amplifier can be designed to operate with very low drift despite changes in ambient temperature, supply voltage, and frequency; typical ranges of variation of temperature −55 to +71 degrees centigrade, supply voltage ±10 per cent of nominal, and supply frequency ±5 per cent of nominal. As the author states, even when using the best available components in the amplifier effective drift minimization requires that the amplifier gain be fixed at a value less than that of the basic self-saturating circuit by means of negative feedback. With reference to Fig. 7, we should be interested in knowing the net feedback ratio at each gain level shown.

The author's definition of an operational amplifier described in the section titled "Basic Theory of Operation" is somewhat restrictive. This definition is limited to the performance of algebraic functions in a system. We believe that a broader and more valuable definition should include the performance of these functions plus operations in the frequency domain. By the addition of frequency sensitive elements in series or in feedback loops it is possible to provide system frequency compensation with such an amplifier. In this way the amplifier imposes its characteristics on the feedback control system with beneficial results.

R. M. Hubbard: I wish to thank Mr. Horton and Mr. Decker for their discussion. Before answering their important question concerning the net feedback ratios associated with the several gain levels shown in Fig. 7, I should like to point out that a straightforward calculation of these ratios involves the determination of several feedbacks: The positive self-saturating feedback produced by the rectifiers in series with the load windings; the positive feedback introduced by the coupling circuit, see Fig. 3; the negative feedback inherent in the bias circuit; the differential positive feedback added to raise the amplifier gain to infinity; and the stabilizing negative feedback which fixes the amplifier gain level. It seems advisable, therefore, to answer this question instead by comparing the resulting net feedback ratios for the curves of Fig. 7 with that of a reference circuit comprising a simple, fixed-bias self-saturating amplifier with no coupling circuit. The ratios of the net feedbacks for curves 1 through 5 of Fig. 7 to that of the reference circuit are as follows:

No. 1,	1.00
No. 2,	0.99
No. 3,	0.98
No. 4,	0.96
No. 5,	0.93

The definition of an operational amplifier is not intended to exclude operations in the frequency domain even though the examples given involve only performance of algebraic functions. The point to be made is that if an amplifier satisfies the requirements for an operational device it can be used to perform such functions as integrations, for example, without imposing its own time constant on the system. Actually the term "operational amplifier" as used in this paper was intended to apply only to the basic infinite-gain device, so that any and all functions that are performed are exactly predictable on the basis of the configuration of the negative feedback loop.

Negative-Impedance Telephone Repeaters—Application in the Bell System

ARTHUR F. ROSE
FELLOW AIEE

THE application of the telephone repeater, the development of which made countrywide telephone service practicable, had been confined largely to toll plant from the year 1915 when the transcontinental line was first established, until a few years ago. About 1948 the negative-impedance repeater[1] was developed and placed in production. This repeater operates on the principle of inserting negative resistance (and, if desired, negative inductance or capacitance) in series with the line, thus reducing the over-all impedance and increasing the current in the line. This results in transmission gain in the same sense as that resulting from a repeater of the conventional type. This principle and the package nature of the assembly resulted in a telephone repeater so low in cost and so simple in application and installation that it has found extensive use primarily in the local telephone plant.

In the period from 1948 to the present, over 50,000 series-type negative-impedance repeaters were manufactured and incorporated in the Bell System telephone plant. These repeaters have been used largely on intraexchange trunks and on trunks extending from the exchange areas to near-by smaller towns. Such installations have been very effective in improving the transmission on short-haul calls and in many cases have also reduced trunk costs by permitting the use of smaller and cheaper conductors. They are usually operated at gains which reduce the nonrepeatered trunk loss by more than half.

Negative-impedance repeaters are especially suited to exchange trunk use, not only because of their simplicity and ease of maintenance but also because, unlike earlier types of repeaters, they preserve the d-c continuity of the circuits on which they are installed. This latter feature means that they do not interfere appreciably with the signaling methods ordinarily used on such trunks. Also, they have the added advantage that in the event of tube failure the circuit still func-

---

Paper 54-222, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 25, 1954; made available for printing April 14, 1954.

A. F. ROSE is with the American Telephone and Telegraph Company, New York, N. Y.

tions but with its loss substantially raised until the defective tube has been replaced.

Application of Series-Type Negative-Impedance Repeaters

In a multioffice exchange area, trunks can generally be grouped in three classes. Largest in number are those known as inter-office trunks which extend directly from one local operating center to another local center in the same exchange operating area. In some cases direct trunks between two centers can not be justified and each office has trunks to a central location known as a tandem center where a through connection is made when required. In this second class the trunks are known as tandem trunks. Frequently trunks of this type are provided to supplement the direct trunks and thus give the added advantage of alternate routing. A third type of trunk extends between the local office and the toll office, and therefore always forms part of a toll connection to the local office. These are known as toll-connecting trunks. In single-office areas the toll connecting trunks to neighboring small towns are known as tributary trunks.

In some large multioffice exchange areas there are also trunks between the toll office and tandem centers which may be some distance from the toll office and through which local offices reach the toll office for long-distance calls. These trunks are also called tandem trunks but are more in the nature of intertoll links and are required to operate at very low loss, similarly to those in an intertoll circuit.[2]

The most extensive use of the negative-impedance repeaters has been in connection with the interoffice type of trunk. In this role they have been extremely useful and have contributed greatly to transmission improvement in the last few years. It would be desirable also to utilize this new and effective tool in reducing the losses of toll-connecting and tandem trunks. There have been many cases of this type where improvement has been effected but a very wide use has been limited by their effect on return loss.

The insertion of a series negative impedance in an otherwise uniform loaded circuit introduces an impedance discontinuity which means that a substantial amount of energy is returned to the sending end of the circuit as "echo." As the introduction of negative impedance is the method of obtaining gain from the series repeater, it follows that the magnitude of the impedance discontinuity and therefore, the echo is proportional to the gain of the repeater. The effect of echo on ease of conversation depends both on its magnitude and the amount of delay[3] before the echo reaches the talker's ear. When the delay is short, as in the case of an interoffice trunk, a large amount can be tolerated but in the case of a toll-connecting trunk which becomes a part of a toll connection, the delay may be long and only a small amount of echo can be permitted. This limits the amount of gain from a series repeater when it is used in toll-connecting trunks.

As a result of this feature of the series negative-impedance repeater, most cases requiring substantial gains in toll-connecting or tributary trunks usually had to be handled by the older and more expensive hybrid-coil type of repeater. With the latter it is practicable to introduce gain without serious reflection effects if proper attention is given to network design and to uniformity of construction of the line itself. However, the cost of the more complicated repeater was relatively high, so that the new development described in the accompanying paper[5] was undertaken to meet the indicated need, i.e., a repeater embodying the same desirable features of simplicity of design but still approaching in performance that of the hybrid-coil repeater in its effect on return loss.

As indicated by Merrill and Smethurst[4,5] the new repeater (coded *E23*) consists essentially of the earlier series repeater with the addition of a shunt negative-impedance element. This combination has approximately as small an effect on return loss as the hybrid-coil repeater and will give about the same gain when used under similar line conditions. As a result, the field of use of the negative-impedance type of repeater will be greatly extended, especially in the toll-connecting plant, where, from a practical standpoint, certain features of the hybrid-coil repeater are not needed. For example, its maximum gain of 10 decibels (db) is adequate, unequal gains in opposite directions would not ordinarily be useful, and types of signaling requiring a 4-wire split of the repeater are not utilized on exchange area trunks.

In addition to the important new characteristics, the desirable features of the series repeater have been retained, i.e., it preserves d-c continuity, it is simple in design, and it is easy to install and maintain. If properly engineered and installed on circuits of reasonably uniform impedance, the new repeaters can be operated in tandem, do not have serious reactions on return loss, and are capable of reducing the losses of the trunks to about the same values as the hybrid-type repeaters.

Scope of Application of the New Repeater

With the reduction of the return loss reaction of the earlier form of negative impedance repeater, the extent of application of the new *E23* repeater becomes largely a matter of economics. In the telephone trunk plant, there are generally three gauges of conductors available for trunk use, namely, 19, 22, and 24. The larger gauges, utilizing more copper and requiring more conduit space, naturally cost more than the smaller. In general, it is the practice to select the smallest gauge which will give the required transmission loss. For short distances, the differential in cost between trunks of dif-

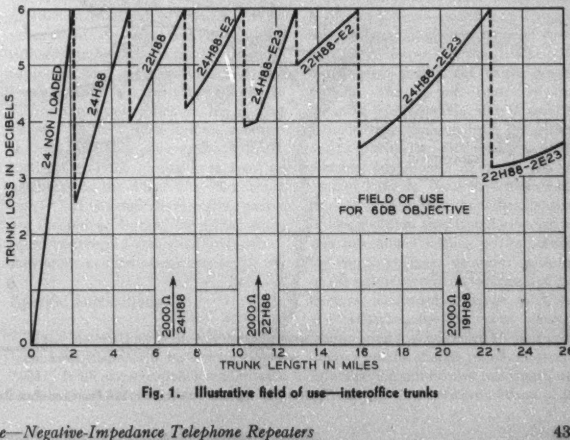


Fig. 1. Illustrative field of use—interoffice trunks

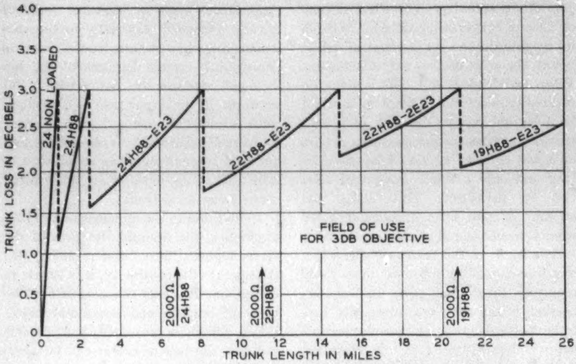


Fig. 2. Illustrative field of use—toll-connecting or tandem trunks

ferent gauges is not large but as trunk length increases it will be found that a small-gauge trunk with a repeater will cost less than one of larger gauge without a repeater. For example, at 10 miles a 22-gauge circuit with a repeater will cost substantially less than a 19-gauge circuit with no repeater. Also, as transmission objectives are improved, there may be cases where a small-gauge trunk with two repeaters will be cheaper than a larger gauge with one repeater. Furthermore, in cases where the costs are about equal or even where the larger gauge is slightly cheaper, it will be found that lower losses can be obtained with the repeatered circuits, and therefore the engineering choice would favor the smaller gauge.

Construction costs vary considerably depending on local conditions. Therefore the differential cost between the various gauges of conductors will vary correspondingly. However, for illustrative purposes two charts have been prepared (Figs. 1 and 2) which utilize average Bell System costs for both outside plant and telephone repeaters. In the case of the *E23* repeaters, the experience so far has been somewhat limited, so there is a greater degree of uncertainty as to the actual costs than for the older type of repeater or the outside plant construction.

Fig. 1 covers the field of use for various conductor gauges when used in interoffice trunks where an over-all transmission objective of six db for the trunks has been assumed. In considering this chart it will be seen that nonloaded 24-gauge conductors can be utilized for the shorter distances up to about 2 miles. The next step is to apply loading, which is indicated in the chart by the symbol *H88* meaning 88-millihenry coils at 6,000 foot spacing. After reaching the limit of about 7 miles without repeaters on 22 gauge, the differential between 22 and 24 gauge is more than enough to pay for a repeater, the simple series repeater *E2* being used for distances up to 10 miles and the improved *E23* repeater extending the use of 24 gauge to about 14 miles. Beyond this point the most economical combination is indicated.

The general effect of the introduction of the negative-impedance repeater is to shift the average gauge distribution so that more small-gauge cable can be economically utilized, accompanied in most

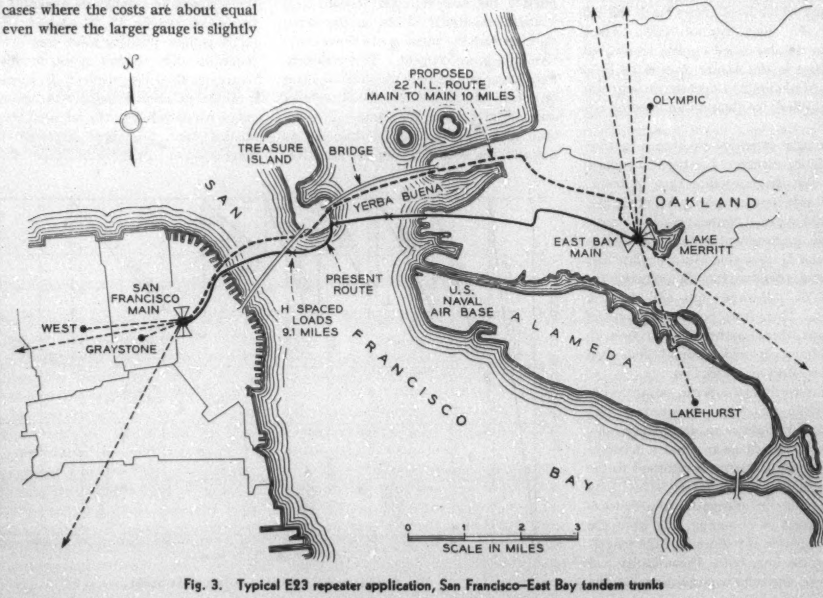
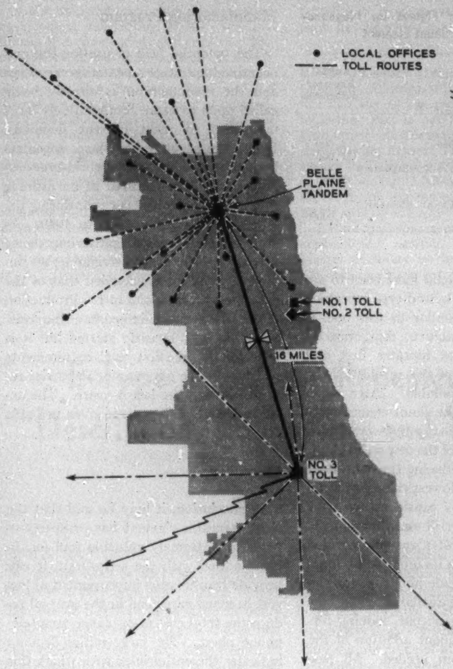


Fig. 3. Typical E23 repeater application, San Francisco–East Bay tandem trunks



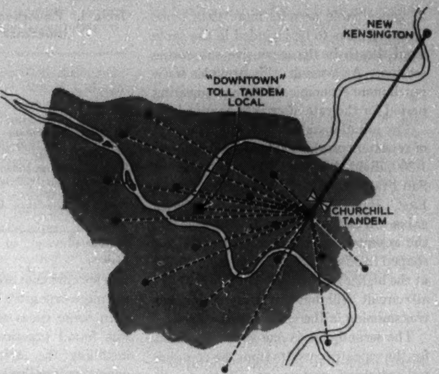


Fig. 4 (left). Typical E23 repeater application, Chicago toll tandem trunks

Fig. 5 (above). Typical E23 repeater application, Pittsburgh toll-connecting trunks

cases by improved transmission. It will be noted from the chart that no 19-gauge cable will need to be added in the future to care for interoffice trunks up to distances as great as 25 miles between offices. Of course, in cases where 19 gauge is already available in plant it can be used to advantage despite the fact that for new construction a smaller gauge with the negative-impedance repeaters would be cheaper.

It has also been assumed in making up this chart that supervision or pulsing requirements will not be effective in limiting the use of the smaller gauges. In a specific case where some of the older type of central offices are involved, signaling may have an important bearing and substantially distort the economic ranges indicated on the chart. Also, the insertion of the repeater has a small effect on the signaling and pulsing ranges but, in general, this is not controlling.

Fig. 2 covers toll-connecting or tandem trunks. As a toll-connecting trunk forms part of a multilink connection and, since there are always at least two (and often more) trunks in series on connections involving these trunks, the Bell System Companies find it economical to plan for a maximum loss of 3 db for this type trunk. Likewise for the tandem trunks, where two in series may be used in place of an interoffice trunk, 3 db is considered a reasonable objective.

It will be noted again by referring to this chart that 19 gauge has a very little future field of use and in all cases the new series-shunt repeater will be utilized rather than the earlier series type because of return loss considerations and also because of the greater gain required to reduce the trunk losses to the desired values. In the tandem and toll-connecting case, signaling may again be a distorting factor though not to as great an extent as in the direct interoffice trunk case. To this extent, however, the curves are theoretical, as they have been made up without regard to this limitation, which may apply in a few practical cases.

Specific Applications

Many installations of the new repeaters have been engineered for completion in 1954. In all cases economic studies were made and results of these studies broadly confirmed the indications of the two charts discussed above. To bring out more clearly the effectiveness of the *E23* repeaters, it may be worth while to consider a few specific cases involving installations of the new repeaters being made this year.

The first case shown in Fig. 3 is of particular local interest as it is in the area of the Pacific Telephone and Telegraph Company. The chart shows the two alternatives were considered for providing additional tandem trunks between San Francisco proper and the East Bay Area, needed this year because of an extension of customer toll-dialing arrangements.

The engineering study for this project was somewhat more complicated than would ordinarily be the case because two possible routes of unequal length and with slightly different amounts of submarine cable were involved. The present route, which touches Yerba Buena Island, is subject to some hazard from dragging ship anchors but the circuit relief would have been cheaper here than on the other route shown, were it not for the new repeaters. The second route is considerably less hazardous and, in addition is sufficiently removed from the present one to provide increased reliability under disaster conditions. Without the repeaters, however, the second route would have been somewhat impracticable, since it does not allow easy installation and access to required loading points between the two shore lines. However, with two of the new repeaters on each pair of conductors, nonloaded 22-gauge cable will give substantially the same effective transmission loss as loaded 19-gauge conductors on the shorter route and over a future period will involve less annual cost per circuit. Furthermore, the use of the

22-gauge cable permits more than twice as many pairs to be included in the same sized sheath in the expensive submarine section. As a result, this provides for a much more economical future engineering period for the rate of growth assumed.

The initial installation of the new type of repeaters on this cable will total about 1,300, divided between the main office in San Francisco and the main office on the East Bay side. Individual repeater gains will be 6 to 7 db at 1,000 cycles but the repeater gain characteristic will be shaped to offset the increasing cable losses at the higher frequencies so that the over-all circuit will have relatively uniform transmission in the voice range.

The second case is one where the need for the repeaters results from the complex toll-switching system at Chicago, Ill. Here it has been found desirable, because of the great volume of toll traffic, to have several toll offices scattered throughout the city and suburbs. In general, toll calls coming into these offices from other cities are completed over toll-connecting trunks direct to the called subscribers' central offices. It is not economical, however, to provide a sufficient number of such trunks to handle peak loads. When all of these direct trunks are busy, an incoming toll call is switched to the subscriber's central office via a tandem office serving his general area. Since the losses of trunks between the tandem and local offices are of the same order as those of the direct trunks from toll to local offices, it would be desirable to operate the trunks between the toll and tandem offices at losses close to 0 db if this were practicable. In this way the same transmission objectives would be met on both routings and no contrast in transmission would be evident to the same subscribers on calls completed over the different routes at different times. Fig. 4 illustrates a specific case of 36 trunks between toll office no. 3 and Belle Plaine tandem, a distance of about 16 miles. Without repeaters these trunks would have been operated at a loss of about 10 db but with the repeaters they are operated at about 2 db. The only alternative to using *E* repeaters in this specific case would have been to use the more expensive hybrid-type repeaters.

Table I. Production Figures for Negative-Impedance Shunt Element

	Annual Repeater Shipments
Hybrid-type repeaters	
Average year 1930–1949	4,500
Maximum year	17,000
Series negative impedance-type repeaters	
Average year 1950–1954	17,000
Maximum year	36,000
Negative Impedance Shunt Element, 1954	19,000*

*Estimated.

In some cases similar to this one, it was found practicable to temporize by installing the series repeater first and operating it at limited gain until the shunt element became available. Later, when the production of the shunt element was started, the additional units were added and full advantage of the new series-shunt design utilized in reducing the equivalent of the trunks to the lowest value permitted by echo return loss considerations. As pointed out by Merrill and Smethurst[4,5] the original *E1* repeater was not provided with a center tap on its input coil but arrangements were made about a year ago to make this feature available for the current manufacturing output, looking forward to the subsequent addition of the shunt element when needed. All new models of the series-type repeater now include the center-tapped coil.

The third example, shown on Fig. 5, is in the city of Pittsburgh, Pa., between the Churchill tandem and the town of New Kensington, approximately 17 miles. Here the transmission on existing 19-gauge loaded cable without a repeater would have been about 8 db. The repeaters reduce this figure to between 2 and 3 db which should be satisfactory in this case. It should be noted, however, that 22 gauge with two repeaters would also provide a low enough equivalent, and should it become necessary to supplement the present cable at a future date there will be an opportunity for further saving from the *E23* repeaters.

Manufacturing Program

The potential field of use for the new negative-impedance repeaters is very large and the manufacturer is already being called upon to make comparatively large quantities to meet current demands. Before the negative-impedance repeaters became available, the hybrid-type repeater was manufactured at an average rate of less than 5,000 per year during the 20-year period from 1930 to 1949. As soon as the series negative-impedance type of repeater became available, the annual demand for it exceeded that of the old type by a ratio of 3 to 1. Production of the new negative-impedance shunt element has only recently started but it is expected that the first year's requirements will exceed the average for the series repeater during the last 4 years. The actual production figures are given in Table I.

Summary

In conclusion, it may be said that the shunt repeater element has prospects of being an extremely valuable tool for, in combination with the series unit, it can provide transmission improvement at low cost in many cases and in the over-all reduce the total cost of exchange area telephone plant. The large initial demand indicates the enthusiasm with which this new development is being greeted by the Telephone Companies throughout the Bell System.

References

1. A NEGATIVE IMPEDANCE REPEATER, **J. L. Merrill, Jr.** *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1461–66.

2. FUNDAMENTAL PLANS FOR TOLL TELEPHONE PLANT, **J. J. Pilliod.** *AIEE Transactions*, vol. 71, pt. I, Sept. 1952, pp. 248–56.

3. THE TIME FACTOR IN TELEPHONE TRANSMISSION, **O. B. Blackwell.** *AIEE Transactions*, vol. 51, March 1932, pp. 141–47.

4. THEORY OF E-TYPE REPEATERS, **J. L. Merrill, Jr.** *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 443–47.

5. E-TYPE TELEPHONE REPEATERS—DESCRIPTION, EQUIPMENT AND TESTING, **J. O. Smethurst.** *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 435–43.

◆

Discussion

D. I. Cone and **F. H. Wright** (The Pacific Telephone and Telegraph Company, San Francisco, Calif.): It is a sincere tribute to the developers of the new negative impedance repeaters to say that the *E-23* repeater has made feasible an important contribution to the adequacy and dependability of the telephone network linking the communities about San Francisco Bay. Referring to Fig. 3 of Mr. Rose's paper shows the heart of the San Francisco Bay metropolitan area, served by about 1,000,000 telephones. The number of trunk circuits between local telephone offices is rapidly approaching 100,000, about 6,000 of which are in the submarine path shown between San Francisco and Oakland.

The new repeaters were made available just in time for our 1954 needs, and they made possible a new route, free of the exposed loading points at 6,000-foot intervals required by most of the existing circuits. One of these loading points is on a Bay Bridge pier, another on a special structure in shallow water. The results of tests of the new repeaters and submarine cable have confirmed the predictions of performance. Effective transmission losses equivalent to *19H88* cable for the same distance have been realized.

It is anticipated that future carrier system development will further increase the usefulness of the nonloaded cable circuits. The existing route has long handled *K* and *N* carrier systems. For television transmission across San Francisco Bay, a *TD-2* microwave link is used. Costs of channelizing and signaling have precluded the use of radio for transbay telephone trunks.

The *E* repeaters have proved so attractive for short-haul trunk circuits that we anticipate about 9,000 in use in 1956 in the Bay area. They have proved very valuable in adapting existing plant to the requirements of private-line services, such as are needed for military and industrial circuits.

A recent case of transmission deficiency on a private line was solved by using a portable *E* repeater set, including power pack, developed for just such contingencies.

R. R. O'Connor (Illinois Bell Telephone Company, Chicago, Ill.): We in Chicago feel that the *E*-type repeaters will play an important part in our trunk plant design. How important this may be is illustrated by the following figures: The total number of trunks in Chicago is 70,000. By the end of 1954, 6,000 (1,300 *E23*) repeaters are expected; by the end of 1955, 9,300 (3,000 *E23*); and by the end of 1960, 30,000.

With *E*-type repeaters we expect to find it economically possible to provide much greater flexibility in the routing of traffic. This will result partly from alternate routing via tandem and partly from the use of separate cable sheaths for a single trunk group. In the latter case one cable will be equipped with repeaters and another of coarser gauge will be used without repeaters.

Arthur F. Rose: The discussions by Mr. Cone and Mr. Wright and by Mr. O'Connor emphasize the rapid adoption of this new repeater in the telephone plant of the Bell System by citing specific figures for the Pacific and Illinois Bell Telephone Companies. The use of the *E* repeater in portable form as mentioned in Mr. Cone's and Mr. Wright's discussion is an example of the versatility of this new tool which it is hoped will continue to find new applications.

E-Type Telephone Repeaters—Description, Equipment, and Testing

J. O. SMETHURST
NONMEMBER AIEE

THE insertion of repeaters in telephone lines to make up for circuit losses is a very common practice today. The use of negative impedance to overcome transmission loss is an old idea but only recently used to any large extent in the exchange area plant. A series-type of negative-impedance repeater, called the *E1* telephone repeater, has found favor in the Bell System and its use has stimulated a demand for still more economical designs including a companion shunt type of negative-impedance repeater. Sufficient laboratory and field trial work has been done on negative-impedance devices to insure that better transmission performance can be obtained, at minimum cost, by using both a shunt element, of the negative-admittance type, and a series type of negative-impedance repeater. Through the use of these new *E*-type repeaters it is possible to reduce the attenuation and improve the return loss of voice frequency lines so that exchange area trunks can be used as links in a toll connection.

Rather than adapting some form of modification to the existing *E1* repeater design, many advantages are obtained by producing a new series-type of negative-impedance device, called the *E2* repeater, and a new shunt type of negative-impedance unit called the *E3* repeater. New equipment designs will reduce the maintenance, obtain a substantial reduction in price and make more efficient use of the central office space, particularly when 23-inch relay rack bays are used. Plug-in type of construction will simplify the installation and testing of these repeaters.

A new portable *E*-repeater test set has been developed for simplifying the testing and maintenance of all types of negative-impedance repeaters. Test connections required to measure individual repeaters or combinations of *E1* or *E2* and *E3* repeaters are set up by the operation of a single-control rotary-function switch. Insertion gain and loss measurements are made with the repeaters working between their normally connected line impedances or specified terminations. Jack-ended *E2* and *E3* repeater networks have been included for the rapid setting up of network strapping arrangements during the testing period.

Other features of the test set include means for measuring return losses and impedance. The magnitude of the unknown impedance and return loss and the phase angle of the impedance can be determined by simple transmission measurements.

E2 Telephone Repeater

The circuit functions of the *E2* repeater can be divided into three main parts, as shown in Fig. 1. These include an input transformer, a vacuum-tube impedance converter and the adjustable impedance network. The essential element of this repeater is the impedance converter which basically is a positive feedback grounded-grid amplifier designed to be open-circuit stable. The positive feedback is stabilized by the application of negative feedback and the proportion of positive to negative feedback is controlled by changing the magnitudes of the plate and cathode load impedances. Where the plate load consists of the network impedance Z_N and the cathode load is the effective line impedance reflected through the transformer into each cathode circuit. Small cathode impedances and large plate impedances will reduce the negative cathode feedback, increase the over-all circuit gain, and provide greater amplified voltages for the positive feedback connection. The reverse conditions of large cathode impedances and small plate impedances result in less gain and reduced potential for the positive feedback action.

The positive feedback loop is joined by connecting the plate circuit of each vacuum tube through a resistance-capacitance network to the grid circuit of the opposite tube. The negative feedback path includes the effective cathode impedance and grid circuit of each triode tube. A phasor addition of these two feedback voltages in the grid circuit can be controlled by means of the network impedances, to produce a stable positive feedback condition for any practical value of cathode impedance.

Any signal voltage impressed across the cathode impedances will be applied to the grids and amplified. The increased signal amplitudes, developed across the network impedance, are coupled back to the proper grid where regenerative action causes amplified voltages to appear

---

Paper 54-287, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 25, 1954; made available for printing April 30, 1954.

J. O. Smethurst is with Bell Telephone Laboratories, Inc., Murray Hill, N. J.

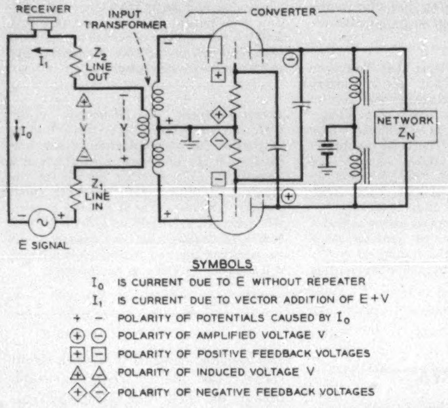

RECEIVER
INPUT TRANSFORMER
CONVERTER
I_1
Z_2 LINE OUT
I_0
V
NETWORK Z_N
Z_1 LINE IN
E SIGNAL
SYMBOLS
I_0 IS CURRENT DUE TO E WITHOUT REPEATER
I_1 IS CURRENT DUE TO VECTOR ADDITION OF E+V
+ − POLARITY OF POTENTIALS CAUSED BY I_0
⊕ ⊖ POLARITY OF AMPLIFIED VOLTAGE V
⊞ ⊟ POLARITY OF POSITIVE FEEDBACK VOLTAGES
▲ △ POLARITY OF INDUCED VOLTAGE V
◈ ◇ POLARITY OF NEGATIVE FEEDBACK VOLTAGES

Fig. 1. Simplified circuit of E2 repeater

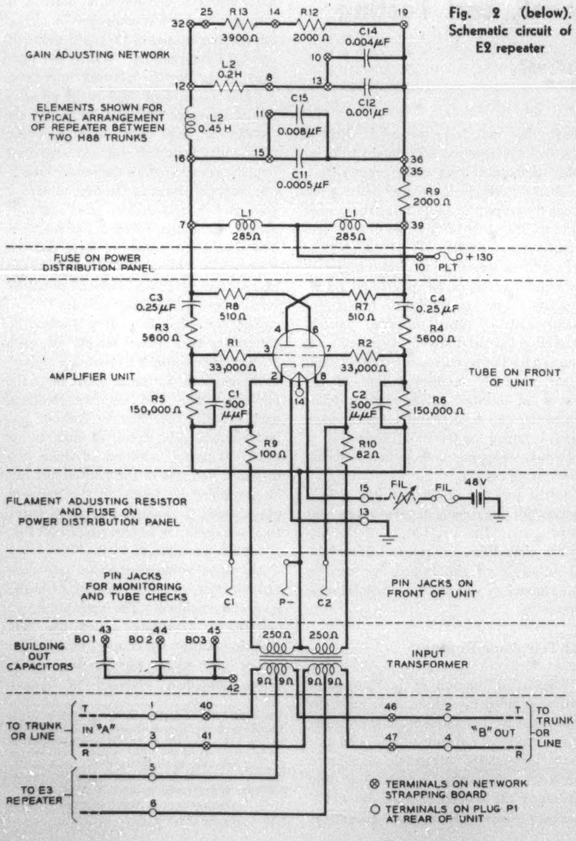

GAIN ADJUSTING NETWORK
R13 3900Ω
R12 2000Ω
C14 0.004μF
L2 0.2H
C12 0.001μF
C15 0.008μF
ELEMENTS SHOWN FOR TYPICAL ARRANGEMENT OF REPEATER BETWEEN TWO H88 TRUNKS
L2 0.45H
C11 0.0005μF
R9 2000Ω
L1 285Ω
L1 285Ω
FUSE ON POWER DISTRIBUTION PANEL
10 PLT +130
C3 0.25μF
R8 510Ω
R7 510Ω
C4 0.25μF
R3 5600Ω
R4 5600Ω
R1 33,000Ω
R2 33,000Ω
AMPLIFIER UNIT
TUBE ON FRONT OF UNIT
R5 150,000Ω
C1 500 μμF
C2 500 μμF
R6 150,000Ω
R9 100Ω
R10 82Ω
FIL FIL 48V
FILAMENT ADJUSTING RESISTOR AND FUSE ON POWER DISTRIBUTION PANEL
PIN JACKS FOR MONITORING AND TUBE CHECKS
C1 P− C2
PIN JACKS ON FRONT OF UNIT
BUILDING OUT CAPACITORS
BO1 BO2 BO3
250Ω 250Ω
9Ω 9Ω 9Ω 9Ω
INPUT TRANSFORMER
TO TRUNK OR LINE
IN "A"
"B" OUT
TO TRUNK OR LINE
TO E3 REPEATER
⊗ TERMINALS ON NETWORK STRAPPING BOARD
○ TERMINALS ON PLUG P1 AT REAR OF UNIT

Fig. 2 (below). Schematic circuit of E2 repeater

across the cathode impedance in phase opposition to the applied signal. A phasor addition of the two cathode voltages will occur in the transformer windings and a large reversed signal potential will be inserted in series with the line. The induced voltage, being of the same polarity as the originating signal, will add to it and cause a larger value of signal current to flow in the line circuit. Because larger line currents are obtained by the use of a 2-terminal network, the *E2* repeater can be considered as having the properties of a negative impedance equal to $-V/I$, where V is the reversed voltage across the impedance and I the current flowing through it.

Merrill[1] has shown that the negative impedance generated by this form of converter is equal to the negative of the network impedance multiplied by a conversion factor k. To reduce the effect of variations in the vacuum tubes to negligible proportions, and at the same time to operate the tubes with load impedances which will permit optimum energy transfer from tube to connected circuit, the converter should be used with an impedance of 10,000 ohms or more between cathodes and a somewhat lower impedance between plates. To obtain an impedance of approximately 10,000 ohms, the impedances Z_1 and Z_2 of Fig. 1 are stepped up to the cathodes by means of the input transformer. In addition, applications of the repeater to telephone lines requires that a balance-to-ground condition be maintained to meet the longitudinal balance and noise requirements. To insure adequate balance the low-voltage side of the transformer is divided into two equal balanced windings. Each winding is center-tapped and connected in series with a line conductor.

The conversion ratio is affected also by small losses inherent to the vacuum tube and transformer circuit designs. These may be balanced out by placing a fixed compensating resistor of 2,000 ohms in the network impedance. In practice it is advantageous to limit the conversion frequency bandwidth so that line and network impedances do not have to be controlled over unlimited frequency range. The conversion of the *E2* repeater is limited at the low frequencies by a grid-to-plate coupling network and at the high frequencies by small capacitances bridged across the grid resistors, as shown in Fig. 2.

The final negative series impedance presented to the line is equal to approximately $-0.1Z_N$ over the frequency band of 300 to 3,500 cycles. The magnitude and phase of the negative impedances are

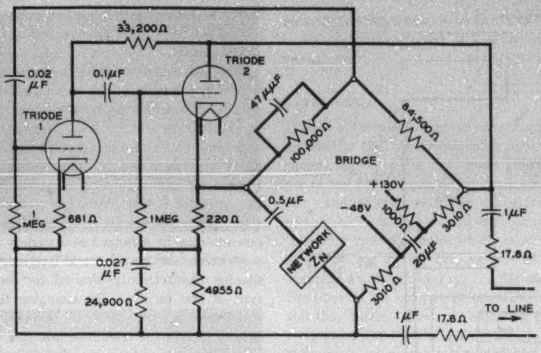

33,200Ω
0.02 µF
0.1µF
TRIODE 2
TRIODE 1
47µµF
100,000Ω
84,500Ω
BRIDGE
+130V
0.5µF
-48V
1000Ω
3010Ω
20µF
3010Ω
1µF
17.8Ω
1 MEG
681Ω
IMEG
220Ω
0.027 µF
24,900Ω
4955Ω
NETWORK Z_N
3010Ω
1µF
17.8Ω
TO LINE

Fig. 3. Circuit of E3 repeater

controlled by the configuration of the gain-adjusting network Z_N, comprising several inductive, capacitive, and resistive elements. These components may be arranged in any form to obtain the gain- and frequency-shaping characteristic desired for each type of line facility. By changing the value of Z_N the negative impedance inserted in series with the line can be varied from -100 to $-2{,}000$ ohms. This range of negative impedances is adequate for the line facilities with which the repeater will be used.

The *E2* repeater employs a *407A* twin-triode vacuum tube of the 9-pin miniature type. When operated from a plate battery potential of 130 volts the repeater will pass speech volumes of $+10$ *vu* before noticeable distortion is caused by overloading of the vacuum tube. The *407A* vacuum-tube heater circuit can be operated from a 24- or 48-volt office battery. The heater current is 100 milliamperes (ma) for 20-volt operation and 50 ma for 40-volt operation, and the plate current is 11 ma.

E3 Telephone Repeater

The *E3* repeater is a 2-terminal device which uses a stable positive feedback amplifier to generate a shunt type of negative impedance. The circuit functions, shown in Fig. 3, include a high-impedance input circuit, the impedance converter, and the impedance network. Of special interest is the converter circuit which consists of a 2-stage vacuum-tube amplifier working through a Wheatstone bridge circuit in which the telephone line and impedance network are involved as two of the ratio arms. Two resistive components complete the bridge circuit, and the amplifier input and output connections, being at equal potential points, are isolated by the effective balance of the bridge. The connections from the line to the amplifier input and from the amplifier output to the line are made at unequal potential points to provide a positive feedback connection through the bridge, amplifier, and line circuits. When the bridge circuit is unbalanced, by making the impedance of the line smaller than the impedance of the network, a portion of the amplified output is coupled back to the grid as a negative feedback voltage. As the line impedance becomes smaller and smaller, the negative feedback increases and so the converter is short-circuit stable. The phasor addition of the positive and negative feedback voltages in the grid circuit can be controlled, by means of the network impedance, to produce a stable positive feedback condition for any practical value of line impedance.

When the repeater is bridged across a telephone trunk, the signal voltage across the circuit will be impressed across the line side of the output bridge with the polarities shown in Fig. 4. This potential will be connected through the bridge to the input grid, amplified by both tubes, and its phase rotated to produce an enlarged output voltage, across the line, in phase with the applied voltage. The new generator, established in the plate circuit of the second-stage amplifier tube, will send current I_R through the network into the line where it divides I_2 flowing toward Z_2 and I_3 toward Z_1. The phasor addition of I_1 and I_2 will produce a larger flow of current in the receiving load Z_2 and the repeater can be considered as having the properties of a shunt-type of negative impedance equal to $V/-I$ where V is the voltage across the impedance and I is the reverse current flowing through it.

It has been shown[1] that the negative impedance generated by this form of converter is equal to the network impedance Z_N divided by a negative conversion factor k. To obtain a practical design of the *E3* repeater for the faithful conversion

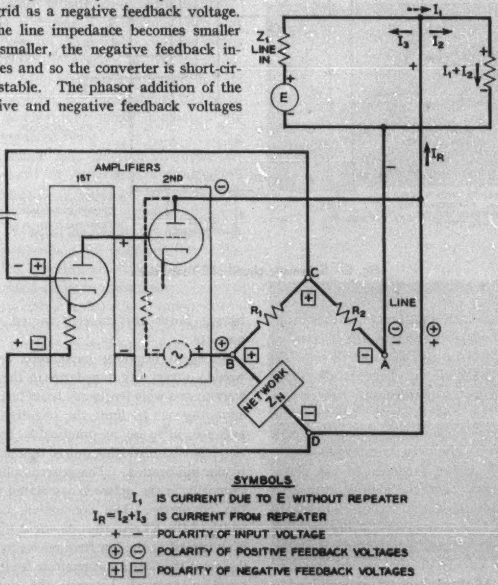

I₁
Z₁ LINE IN
I₃
I₂
I₁+I₂
Z₂ LINE OUT
E
I_R
AMPLIFIERS
1ST
2ND
C
R₁
R₂
LINE
B
A
NETWORK Z_N
D
SYMBOLS
I₁ IS CURRENT DUE TO E WITHOUT REPEATER
I_R=I₂+I₃ IS CURRENT FROM REPEATER
+ − POLARITY OF INPUT VOLTAGE
⊕ ⊖ POLARITY OF POSITIVE FEEDBACK VOLTAGES
⊞ ⊟ POLARITY OF NEGATIVE FEEDBACK VOLTAGES

Fig. 4. Simplified circuit of E3 repeater

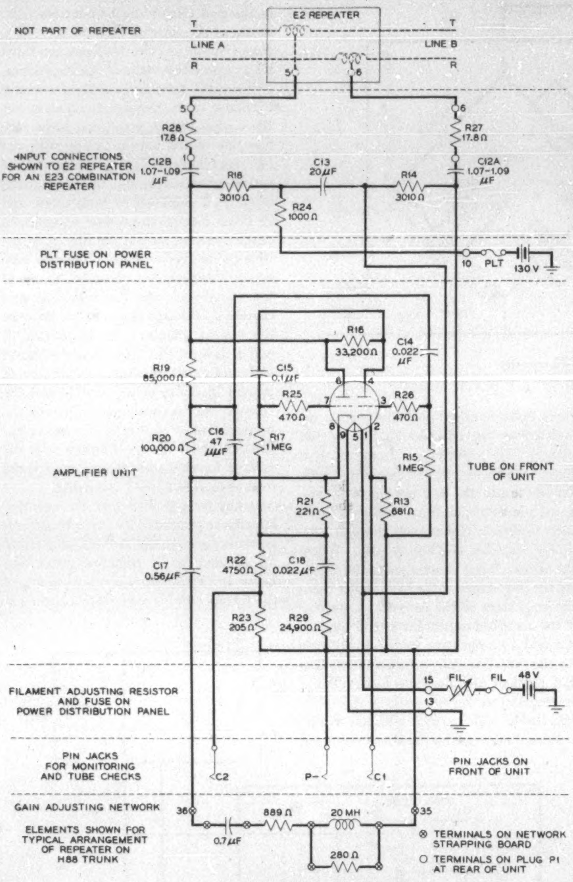

NOT PART OF REPEATER
E2 REPEATER
T
T
LINE A
LINE B
R
R
5
6
5
6
R28
17.8 Ω
R27
17.8 Ω
INPUT CONNECTIONS
SHOWN TO E2 REPEATER
FOR AN E23 COMBINATION
REPEATER
1
C12B
1.07–1.09
μF
C13
20μF
C12A
1.07–1.09
μF
R16
3010 Ω
R14
3010 Ω
R24
1000 Ω
PLT FUSE ON POWER
DISTRIBUTION PANEL
10
PLT
130 V
R18
33,200 Ω
C14
0.022
μF
R19
85,000 Ω
C15
0.1μF
6
4
R25
470 Ω
7
3
R26
470 Ω
8
2
9
5
1
R20
100,000 Ω
C16
47
μμF
R17
1 MEG
R15
1 MEG
AMPLIFIER UNIT
TUBE ON FRONT
OF UNIT
R21
221 Ω
R13
681 Ω
C17
0.56μF
R22
4750 Ω
C18
0.022μF
R23
205 Ω
R29
24,900 Ω
FILAMENT ADJUSTING RESISTOR
AND FUSE ON
POWER DISTRIBUTION PANEL
15
FIL
FIL
48 V
13
PIN JACKS
FOR MONITORING
AND TUBE CHECKS
C2
P–
C1
PIN JACKS ON
FRONT OF UNIT
GAIN ADJUSTING NETWORK
36
35
889 Ω
20 MH
0.7μF
280 Ω
ELEMENTS SHOWN FOR
TYPICAL ARRANGEMENT
OF REPEATER ON
H88 TRUNK
⊗ TERMINALS ON NETWORK
STRAPPING BOARD
○ TERMINALS ON PLUG P1
AT REAR OF UNIT

Fig. 5. Schematic circuit of E3 repeater

of the network impedances, with a minimum of spurious components, it is necessary to balance out, as nearly as possible, all converted circuit elements associated with the output bridge and connections to the line. Accordingly, the two line capacitors are balanced out by a network capacitor. The battery supply resistors, the resistor and capacitor of the plate-battery filter, and the plate-load resistor of the first-amplifier stage are all balanced out in the network by placing suitable values of resistors and capacitors in the cathode circuit of the second-amplifier stage. All these elements are combined to form an equivalent 2-terminal network having fewer and smaller valued components.

An ideal negative impedance device would convert any impedance in the network over a wide frequency band but it is advantageous to limit the negative impedance, in so far as practicable, to the frequency bandwidth required by the particular application. This is accomplished primarily in the network associated with the converter. The conversion bandwidth of the *E3* repeater, shown in Fig. 5, is restricted at the low frequencies by the design of the resistive-capacitive feedback coupling network between the output bridge and the input grid, and at the high frequencies by the shunt capacitance connected across one resistive arm of the bridge circuit.

The final negative impedance shunted across the line is equal to $-Z_N/0.94$, within ± 2.5 per cent, over the frequency range of 200 to 5,000 cycles. The magnitude and phase of the negative impedances are controlled by the configuration of the gain-adjusting network Z_N, consisting of several inductive, capacitive, and resistive elements. These components may be arranged in a variety of ways to obtain the gain- and frequency-shaping characteristics desired for each type of line facility. By changing the value of Z_N, the negative impedance shunted across the telephone line can be varied between -200 and $-3{,}000$ ohms. This range of negative impedances is adequate for the line facilities with which the repeater will be used.

The *E3* repeater employs a *407A* twin-triode vacuum tube of the 9-pin miniature type. When operated from a plate-battery potential of 150 volts the repeater will pass speech volumes of $+6$ *vu* before distortion is noticeable because of overloading in the vacuum tubes. The *407A* vacuum-tube heater circuit can be operated from 24 or 48-volt office battery. The heater current is 100 ma for 20-volt operation, 50 ma for 40-volt operation and the plate current is 5 ma.

Equipment

General

The equipment design of the *E2* and *E3* repeaters was directed toward the objectives of producing negative-impedance repeaters which would be simpler to manufacture, easier to engineer, install, and maintain, and which would make more efficient use of the mounting space, particularly on 23-inch bays, than the present *E1* telephone repeater. Because of the large quantities involved, over-all savings in manufacturing costs can be realized by using a compact aluminum die-cast shell to house the repeater components, which is designed to utilize fully the mounting space available in depth as well as in breadth and height. Such an arrangement will facilitate manufacture in that it lends itself to the use of parallel thermoplastic strips for the mounting of the "pigtail" type of components.[2] Further savings in shop costs were obtained by co-ordinating the designs of the *E2* and *E3* repeaters for maximum interchangeability of parts. These arrangements should include mechanical details as well as electrical components in so far as the circuit designs will permit.

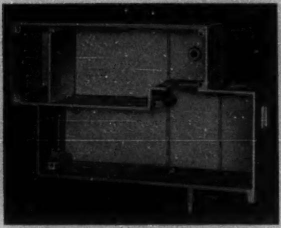


Fig. 6. Die-cast chassis for E repeaters

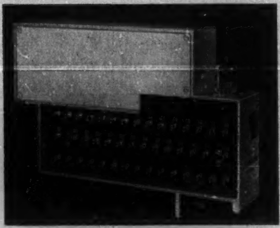


Fig. 7. Front and side view of E2 repeater

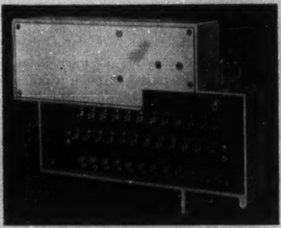


Fig. 8. Front and side view of E3 repeater

Engineering and installation effort has been reduced considerably by avoiding engineered options and by arranging the equipment so that the maximum portion of the assembly and wiring work is performed in the shop. Testing and maintenance routines are simplified by arranging the repeaters as plug-in units which can be removed from their bay positions and plugged into a portable test set located in a more convenient working space. In this way the network strapping and any repair work which may be required on a repeater need not be performed from a ladder. Maintenance and service interruptions are reduced to a minimum length of time by using the plug-in type of repeaters so that a defective unit can be replaced by a spare for immediate restoration of service and the faulty repeater repaired at a more convenient time.

REPEATERS

Both *E2* and *E3* repeater units are assembled in the same rectangular die-cast chassis, shown in Fig. 6, to make efficient use of the full 10-inch depth available in the standard relay rack bay. The front section of each unit carries the vacuum tube and test pin jacks which must be accessible for testing and routine maintenance, as shown in Fig. 7. All other components are compactly assembled inside the chassis. The gain-adjusting network strapping terminals are arranged, in three rows, along the left side of the repeater, as shown in Fig. 8, and are accessible only after the unit is removed from its mounting shelf.

All external connections to either type of repeater are terminated in a male connector rigidly mounted into the repeater chassis. The matching female sockets are suspended from the mounting shelves, on a floating assembly to relieve the strain on contacts and wiring when the repeaters are plugged into a shelf. Both connector units consist of molded rectangular phenolic blocks equipped with 15 gold-plated contacts. Proper alignment between the male and female connectors is maintained by using a positioning key and guide pin on the repeater chassis, shown in Fig. 9, and a track on the mounting shelf. After a repeater is seated into its shelf position, it is made secure by means of a screw-driver operated quick-acting fastener.

A simple and effective mounting arrangement for the large number of pigtail components is obtained by using two parallel thermoplastic strips, as shown in Fig. 10, with the apparatus arranged in two rows along the edges of the strips. Molded brackets are used to fasten the assembled strips to the chassis.

That other apparatus components of various types and sizes may be mounted, with their terminals and leads, in the proper position for wiring and assembly, as shown in Fig. 11, the chassis construction must provide several cells, illustrated in Fig. 6, to contain the equipment. Chassis construction such as this cannot be fabricated economically, even in large quantities, because of the number of operations required. However, they can be designed for economical die casting and in addition to offer a number of other advantages. For example, the die-castings are uniform in dimensions, the cast surfaces are smooth, and the natural aluminum finish has a satisfactory appearance

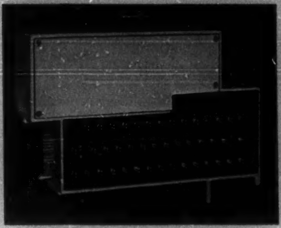


Fig. 9. Side and rear view of E2 repeater

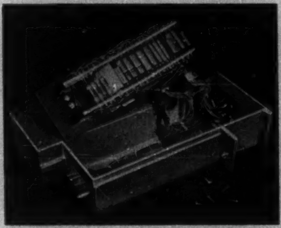


Fig. 10. Network assembly and wiring of E3 repeater

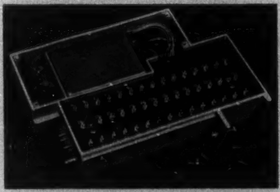


Fig. 11. Assembly and wiring of E2 repeater

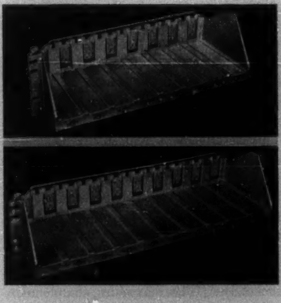


Fig. 12. Repeater mounting shelves. Upper for 19-inch bays, lower for 23-inch bays

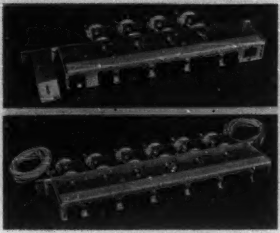


Fig. 13. Power distribution panels. Upper is primary unit, lower is supplementary unit

so that no further finishing operations are necessary.

MOUNTING SHELF

An aluminum die-cast shelf, shown in Fig. 12, is used for mounting the repeaters on the relay rack bays. The shelf comes in two widths, one holding 8 repeaters is used for 19-inch bays and the second containing 10 repeaters is used on 23-inch bays. Grooves cut into the base of the shelf match a projection on the repeater to act as a track system for positioning the repeater into its connector socket. The shelf connector is mounted in a small die-cast block which, in turn, is fastened to the shelf by shoulder screws to provide a floating assembly. A tapered key, molded into the repeater casting, engages a slot in the connector block to secure the horizontal alignment and a tapered pin at the bottom of the repeater casting raises the connector block for the vertical alignment.

The tapered pin is also used for operating micro switches to obtain an automatic line-to-line direct current connection whenever an *E2* repeater is removed from its shelf position. The micro switches, one in each side of the line circuit, are mounted along a ledge on the back of the shelf. A mechanical adjustment is provided to insure a make-before-break connection either from line to line through the switch, when a repeater is removed from the shelf or from line to line through the input transformer of the *E2* repeater when it is installed on the mounting shelf. This arrangement will insure that d-c equipment, such as relays and senders associated with the truck circuits, will not be operated falsely whenever an *E2* repeater is removed for a trouble condition or for routine maintenance.

Fifteen mounting shelves can be arranged on standard 11-foot 6-inch relay rack bays. The maximum complement of repeaters will be 120 for 19-inch bays and 150 for 23-inch bays. Installation costs have been reduced by having the shop install all possible apparatus and wiring on the mounting shelves before shipment to the field.

POWER DISTRIBUTION PANELS

Fabricated power distribution panels which occupy $1^{3}/_{4}$ inches of mounting space are used for supplying plate and filament power to the repeater shelves. The primary unit, shown in Fig. 13, is required for furnishing power to the first three repeater shelves and one test-set power outlet. Four sets of plate and filament alarm-type fuses are provided, one set for each shelf and one set for the test power outlet. Four filament-adjusting rheostats are furnished and the panel is equipped with an alarm relay and lamp. Pin jacks are available for using a voltmeter to measure the filament voltage applied to the vacuum-tube heater circuits. Optional wiring is used for operation of 24- or 48-volt office batteries.

Supplementary power distribution panels are equipped with plate and filament fuses, rheostats and pin jacks for six additional repeater mounting shelves. One primary and two supplementary power panels are required to equip fully an 11-foot 6-inch bay with 15 repeater shelves. Both panels are completely wired in the shop to simplify installation. Fig. 14 shows a 19-inch relay bay mounting shelf complete with *E2* and *E3* repeaters and the two types of power distribution panels.

ALARM-TYPE FUSE

A new alarm-type fuse, coded as the *18A* fuse block, has been used on the power distribution panels for indicating plate and filament fuse alarm conditions. The fuse block, shown in Fig. 15, is completely enclosed and has an added safety feature in that no battery potential is exposed to the attendent when installing or replacing a fuse.

Two alarm features have been built into the fuse and fuse block assembly. The *70*-type fuse used with this block has a spring-loaded tip which pops out through a hole in the fuse block cap, to give a visual indication of a blown fuse. A metal cap, on the end of a blown fuse, makes a connection, through the cap of the fuse block, to an alarm bus which supplies battery for operating the alarm relay and lamp. Three connections protrude through to the rear of the fuse block. One at the bottom, designated *B*, is connected to the power source, one in the center is the fused output and the top contact, designated *A*, is the alarm bus. A small colored pin is inserted into a hole, on the front of the fuse block, to signify the ampere capacity of the fuse required in each particular circuit.

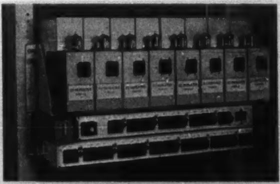


Fig. 14. Shelf of E2 and E3 repeaters, with power distribution panels, for 19-inch bays

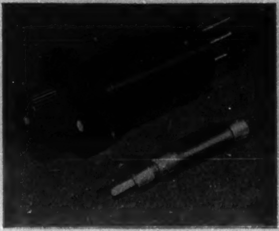


Fig. 15. Alarm fuse block and fuse

All plate and filament fuse-alarm busses are multiplied together in independent paths which are joined through a potential divider circuit for operating the alarm relay and lamp. When a fuse is blown the battery connection operates the alarm relay and lights the alarm lamp on the power distribution panel. The relay supplies the required leads for operating the office visual and audible alarms, which identify the aisle where the failure has occurred; the alarm lamp indicates the bay position; and the telltale protruding tip will locate the actual fuse.

E-Repeater Test Set

Wide applications of the *E2* and *E3* repeaters, to the telephone plant, has made it necessary to develop a new test set which will simplify the installation and maintenance of these devices. To meet this need, the *E*-Repeater Test Set has been designed not only to test the new *E2* and *E3* repeaters but also the older type of *E1* repeater which has been in use for several years. All test connections which are required to measure individual repeaters or combinations of *E1* or *E2* and *E3* repeaters are made by the operation of a single-control rotary-function switch. Such a system will avoid the unintentional errors and time-consuming operations which attend the setting up of complicated patch cord connections. The resulting simplification will reduce the time required for installation and maintenance of these repeaters.

The first five positions of the function switch are arranged to make transmission measurements of an individual repeater or any combination of *E1* or *E2* and *E3* repeaters. Insertion gain and loss measurements are performed with the repeaters working between their normally connected line impedances. In making transmission and stability measurements, it is sometimes necessary to set up trial connections of the repeater networks during the initial test period or for unusual line conditions. The actual use of the repeater networks would require a large number of soldered connections to be made for each particular arrangement of the gain-adjusting network. Through the use of equivalent jack-ended *E2* and *E3* networks, which have been included in the test set, small patch cords may be used for rapid interconnection of the network components.

While the test set was designed principally to test and maintain the *E*-type repeaters, jacks have been added to make the set more adaptable for general field use. With the jack-ended connections, it is possible to measure the insertion gain or loss of other 4-terminal networks in other lines or between specified impedance terminations.

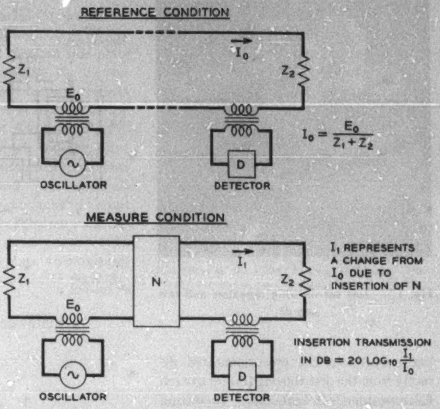
Connections to the test set are made through jacks and connectors into which are plugged the *E2* and *E3* repeaters, the connecting lines from the repeater mounting shelf, and the power cord from the power distribution panel. Connections to the *E1* repeater are made through its own special test cord and connector. An external oscillator and a sensitive detector are required to complete the measuring setup.

THEORY OF TRANSMISSION MEASUREMENTS

The principal application of the *E*-repeater test set is the determination of the insertion gain and stability of the *E*-type repeaters under actual operating conditions. Since gain and stability are dependent upon the line impedances, in both idle and busy conditions, it is necessary to make all measurements with the repeaters connected into their associated lines.

One method of making this type of measurement, without affecting the transmission or stability of the circuit, is to introduce the test voltage through a low-impedance source in series with the input line, and measure the resultant current picked off through a similar low-impedance connection in series with the output line. A comparison of the received currents, with and without a repeater, will give an indication of the insertion gain or loss.

Fig. 16. Principles of insertion measurements


REFERENCE CONDITION
I_0
Z_1
Z_2
E_0
$I_0 = \frac{E_0}{Z_1 + Z_2}$
D
OSCILLATOR
DETECTOR
MEASURE CONDITION
I_1
Z_1
N
Z_2
E_0
I_1 REPRESENTS A CHANGE FROM I_0 DUE TO INSERTION OF N
INSERTION TRANSMISSION IN DB = 20 $\log_{10} \frac{I_1}{I_0}$
D
OSCILLATOR
DETECTOR

The basic principles involved in making this type of transmission measurement can be understood by the examination of the diagrams of Fig. 16. The voltage source and current-measuring detector are assumed to have negligible impedances, compared to the impedances of the lines, a condition which is closely approached in the actual test set.

In the reference condition of Fig. 16, the driving voltage E_0 and the current-measuring device are adjacent and in series with the two line impedances Z_1 and Z_2. The resultant current will be determined by the phasor sum of impedances Z_1 and Z_2 and the source voltage E_0. The current which flows in this circuit may be expressed as follows

$$I_0 = \frac{E_0}{Z_1 + Z_2}$$

where $Z_1 + Z_2$ indicates a phasor addition of the two impedances.

In the measure condition of Fig. 16 the network N is inserted into the circuit between the voltage source and the measuring device and, although the voltage E_0 is assumed to remain constant, the addition of network N into the circuit will change the current to a new value I_1. The change in value of current from I_0 to I_1 is a direct measure of transmission between Z_1 and Z_2 as a result of inserting N into the circuit. The decibels change in transmission caused by the insertion of N can be found from the expression

$$\text{Insertion transmission in } db = 20 \log_{10} \frac{I_1}{I_0}$$

When I_1 is less than I_0 the addition of N has caused an insertion loss and when I_1 is greater than I_0 the addition of N has caused an insertion gain. In practice the current-indicating device will have a decibel scale, and the change in transmission from the reference to the measure condition, can be read directly from the meter scale.

The circuit configuration represented by N may be of any form, such as a series element, a shunt element, combinations of series and shunt elements, sections of transmission lines, or repeaters and amplifiers. The only requirement is that the change in transmission must take place between the point of application of the driving voltage and the point of insertion of the current-measuring device.

The impedances represented by Z_1 and Z_2 may be resistive terminations in order to measure the transmission characteristics of any 4-terminal network, either active or passive, between specified impedance terminations. The only requirement is that the terminating impedances Z_1 and Z_2 must be large compared to the series impedances which are used for insertion of the driving voltage and the detection of the received current.

TRANSMISSION MEASUREMENTS

The transmission section of the *E*-repeater test set is arranged for making insertion transmission measurements, in particular, those measurements required to test and maintain the *E2* and *E3* repeaters in a straight-forward and simple manner. Since the *E2* and *E3* repeaters are of the plug-in type, line and equipment jacks are dispensed with and the

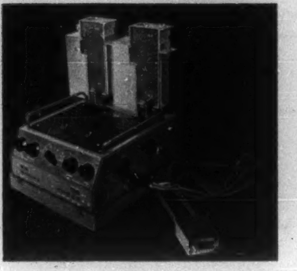


Fig. 17. **Test set showing repeaters and test cords**

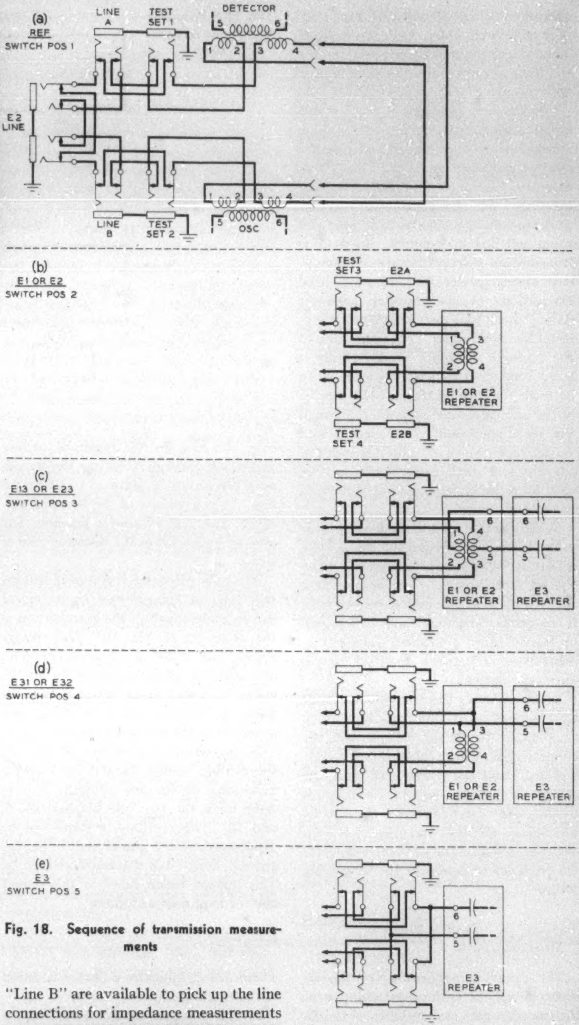


Fig. 18. **Sequence of transmission measurements**

line connections are cross-connected directly from the distributing frames to each *E2* repeater shelf connector. In setting up to make repeater tests, the *E2* and *E3* units are removed from their shelf positions and plugged into adapters in the test set, as shown in Fig. 17.

A test plug, fashioned to simulate the male connector of the repeater, is inserted into the vacant *E2* repeater position for picking up the connections to the incoming and outgoing lines. A jack-ended cord connected to the test plug is patched into the test set for applying the line connections to the repeaters under test. After oscillator, detector, and power cords have been attached to their respective positions the network strapping is set up, on a patch basis, and the test procedures can be performed without further change in the external connections.

The first position of the rotary function switch designated *REF* arranges the test set for the reference condition. All repeaters or networks are removed from the circuit; the sending and receiving series impedances are connected together, and the incoming and outgoing *E2* repeater lines provide the terminations, as shown in Fig. 18(A). The low impedance voltage source is obtained by means of an oscillator working through a step-down transformer having an impedance ratio of 600:2 ohms. The low-impedance side consists of two equal well-balanced windings, one winding inserted into each side of the line to form a balanced-to-ground measuring set. The received current is measured by means of a detector working through an identical transformer having one of the 2-ohm windings connected into each side of the line for maintaining the balance-to-ground circuit. Jacks designated "Test Set 1" and "Test Set 2" are used when transmission measurements are required between specified impedances. Jacks labeled "Line A" and "Line B" are available to pick up the line connections for impedance measurements or for making transmission studies of the lines between fixed impedance terminations. The reference position serves for all four transmission-measuring conditions.

Switch position 2, designated *E1–E2*, connects the *E1* or *E2* repeater into the test circuit between the sending and receiving impedances as shown in Fig. 18(B). The change in detector reading will indicate the insertion gain or loss of the connected repeater. Rapid switching between the reference and measure positions enables the operator to make instantaneous checks of all measurements.

The third and fourth switch positions, labeled *E13–E23* and *E31–E32* respectively, change the arrangements of *E1*, *E2* and *E3* repeaters for the insertion measurements of the various repeater combinations. Figs. 18(C) and (D) show the test connections for *E13* or *E23* repeaters. Fig. 18(E) indicates the connections for testing the *E3* repeater alone on switch position five.

EQUIPMENT

The *E*-repeater test set is arranged as a portable box-type unit 12 by 15 by $8^{1}/_{2}$ inches. A sloping front panel is used with the function switch located in the center with two sections of the resistance standard arranged on each side for ease of manipulation. The main jack field is mounted on a vertical front plate, just below the sloping section, for accessibility and to keep the patch cords away from the controls. Receptacles for the *E1* repeater test cord and power connector are mounted on the right vertical side-plate along with jacks to receive the *E1* repeater center tap connections and connections from the *E2* repeater line. The top plate contains *E2* and *E3* repeater gain-adjusting networks which use the same components and numbering arrangements found in the repeaters. All terminals have been replaced by pin jacks to permit easier interconnection of the network elements by means of small patch cords. At the rear of the top plate, two repeater sockets are mounted to receive the plug-in type of repeaters during the test period. A thumbscrew locking arrangement is used for holding the repeaters securely to the test set. A small compartment has been built into the rear of the set to provide space for storing the network patching cords.

References

1. THEORY OF E-TYPE REPEATERS, J. L. Merrill, Jr. *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 443–47.

2. A NEW TELEPHONE CARRIER SYSTEM FOR MEDIUM-HAUL CIRCUITS, R. S. Caruthers, H. R. Huntley, W. E. Kahl, Ludwig Pedersen. *Electrical Engineering*, vol. 70, Aug. 1951, pp. 692–97.

No Discussion

Theory of E-Type Repeaters

J. L. MERRILL, JR.
NONMEMBER AIEE

IN THE exchange telephone plant, speech is transmitted largely at voice frequencies over a single pair of wires which carries both directions of conversation. Until recently, these circuits were mostly unrepeatered, adequate transmission being assured through a suitable choice of coil loading and conductor size. The need for low-cost repeaters for the exchange plant had long been recognized, but it had not been economical to use widely the conventional hybrid type of repeater with separate amplifiers for the two directions of conversation. The *E1* repeater introduced in 1949 was the first device designed specifically to meet the needs of the exchange plant. It is a bilateral device, which provides gain in both directions of conversation without the use of hybrid coils or line filters and with the use of but a single transmission network.[1] This simplicity, together with the fact that it imposes little impairment on the signaling of the exchange plant, makes it well suited to introducing amplification in 2-wire circuits.

The *E1* repeater has fully met its design objectives but its use introduces an impedance irregularity which limits its application on circuits associated with long-distance connections. Two new repeaters, the *E2* and *E3*, are being introduced this year to meet the requirements of such circuits. The *E2* is electrically the same as the *E1* and differs only in the equipment arrangements, but the *E3* is a new device having no previous commercial counterpart. This repeater is designed for operation in conjunction with the *E2*, and when so used provides an arrangement which can be designed to match the line impedance and thus avoid the undesirable irregularities introduced by a single repeater of the *E1* or *E2* type. The combination *E2–E3* repeater is also a bilateral device which introduces gain in both directions of transmission on a 2-wire circuit without the use of hybrid coils. A convenient way to think of such a device is as an artificial transmission line having a characteristic impedance which matches the impedance of the real circuits with which it is associated but having a negative attenuation constant which replaces some of the energy lost in the real line. The manner in which this is accomplished is the subject of this paper. Equipment arrangements and the application of these repeaters in the Bell System are described in separate papers by Smethurst and Rose.[2,3]

Negative Impedance Concept

The *E2* and *E3* repeater elements each contains an amplifier having multiple feedback paths. The operation of an amplifier circuit of this type can be explained by classical feedback theory. However, experience with the *E1* repeater over the past 4 years has shown the value of using a negative-impedance concept in engineering such a device. Hence, in the explanation of operation given here, the repeater units will be treated simply as 2-terminal networks which have negative-impedance inputs over the frequeny band of interest. The effect of introducing these impedances into telephone circuits can then be computed by the same simple network theory used to determine the effects of passive impedances.

THE E2 REPEATER UNIT

The *E2* repeater is essentially a 2-terminal network, the impedance of which has a magnitude $|Z|$ and a negative phase angle that can vary with increasing frequency from minus 90 degrees, or less, through minus 180 degrees to at least minus 270 degrees. This type of negative impedance is shown in the diagram of Fig. 1(A). For many years it has been known as the series type because it could be produced by connecting the output of an amplifier back in series with its input. More recently it has come to be known as an open-circuit-stable negative impedance because it will not oscillate when its two terminals lie open.

THE E3 REPEATER UNIT

The *E3* repeater is essentially a 2-terminal network the impedance of which has a magnitude $|Z|$ and a positive phase angle that can vary with increasing frequency from plus 90 degrees, or less, through plus 180 degrees to at least plus 270 degrees. This type of negative impedance is shown in Fig. 1(B). It has been known as the shunt type because it could be produced by connecting the output of an amplifier back in shunt with its input. In more recent years it has be-

---

Paper 54-286, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 25, 1954; made available for printing April 30, 1954.

J. L. MERRILL, JR., is with the Bell Telephone Laboratories, Inc., New York, N. Y.

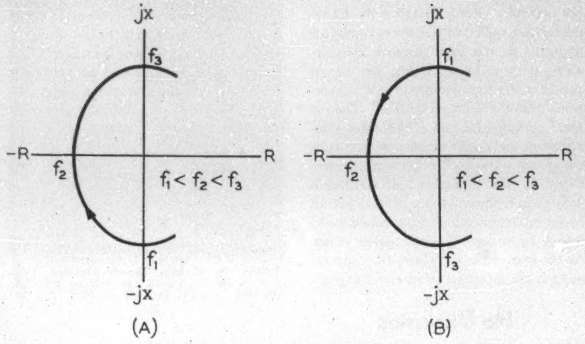


Fig. 1. The two types of negative impedance

A—Open-circuit stable

B—Short-circuit stable

come known as the short-circuit stable type because it will not oscillate when its two terminals are short-circuited together.

THE NEGATIVE-IMPEDANCE CONVERTER

The amplifier circuits of both of these repeater units perform the same function, that of a negative-impedance converter. The operation of such converters is illustrated in Fig. 2.

Fig. 2(A) shows the converter as a 4-terminal network having a ratio of transformation k and a shift of phase through a negative angle of approximately 180 degrees over the operating band of frequencies. If, as shown, an impedance Z_N is connected to terminals 3 and 4 then the impedance seen at terminals 1 and 2 will be the impedance Z_N multiplied by the ratio k and shifted in phase through a negative angle of 180 degrees. This impedance will (over the frequency range of zero to infinity) fulfill the definition of the impedance presented by the *E2* repeater. Hence, Fig. 2(A) can represent the operation of the *E2* repeater.

Fig. 2(B) shows the same converter, but here the impedance Z_N is connected to terminals 1 and 2. The impedance seen at terminals 3 and 4 (at least over the frequency band of interest) will be Z_N divided by k and shifted in phase through a positive angle of approximately 180 degrees. This impedance will (if frequencies from zero to infinity are considered) fulfill the definition given for the *E3* repeater impedance. Thus Fig. 2(B) can represent the operation of the *E3* repeater.

From Fig. 2 it is apparent that the same converter circuit could have been used for both the *E2* and the *E3* repeaters. For practical reasons it was not. However, the ratio k and the phase shift in both the converter of the *E2* and that of the *E3* were made approximately the same. The ratio k was made approximately unity (about 0.94) and the phase shift 180 (plus or minus 15) degrees, negative angle over the frequency range 400 to 3,000 cycles per second.

Operation in Transmission Lines

Within limitations, the *E2* repeater can be represented by a negative impedance $-Z$ and the *E3* repeater can be represented by a negative admittance $-Y$. With a negative impedance and a negative admittance available, losses of transmission lines can be reduced in the manner illustrated in Fig. 3. The transmission line is represented by two networks as shown in Fig. 3(A). One of these (network A) is in the form of a T network, the series arms of which are represented by impedances Z; and the shunt arm by an admittance Y. This network has a propagation constant $\alpha_1 + j\beta_1$. The attenuation α_1 represents the major portion of the line attenuation, and the phase shift β_1 is that just sufficient to make network A realizable physically. This representation is necessary because network A has image impedances both equal to the characteristic impedance (Z_0) of the line. If the characteristic impedance of the line were a pure resist-

Fig. 2 (below). The negative impedance converter

A—E2, Open-circuit stable

B—E3, Short-circuit stable

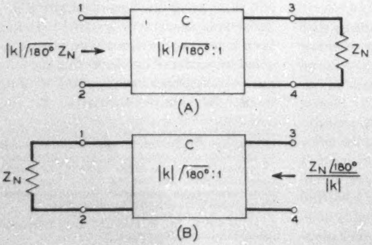
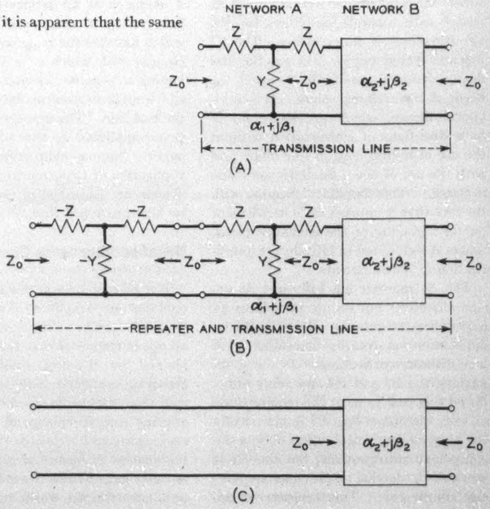


Fig. 3 (right). Operation of the T in a transmission line

A—Transmission line

B—Repeater and transmission line

C—Result of addition of repeater



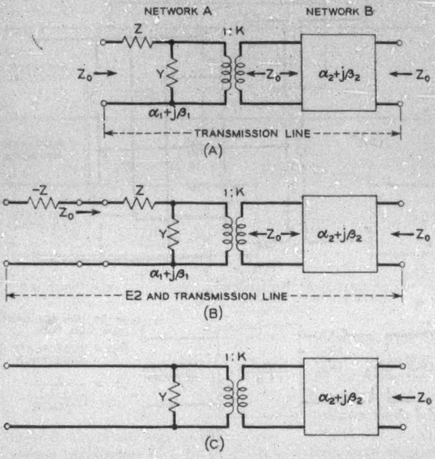

NETWORK A
NETWORK B
Z
1:K
Z_0
Y
$\alpha_2+j\beta_2$
$\alpha_1+j\beta_1$
TRANSMISSION LINE
(A)
−Z
E2 AND TRANSMISSION LINE
(B)
(C)

Fig. 4 (left). Operation of −Z in a transmission line

A—Transmission line
B—Repeater and transmission line
C—Result of adding repeater

ance, then the phase shift through this network could be zero and β_1 could be zero. But the characteristic impedances of actual lines are not pure resistance; thus the phase shift β_1 must be included in network *A*. The other network (*B*) is shown as a box. It has a propagation constant $\alpha_2+j\beta_2$. Here β_2 represents the remaining phase shift in the transmission line, and α_2 is an attenuation just sufficient to make network *B* physically realizable in view of the image impedances which are both equal to Z_0, the characteristic impedance of the line. Fig. 3(B) shows the addition to this line of a repeater consisting of a *T* network made up of negative impedances $-Z$ in the series arms and a negative admittance $-Y$ in the shunt arm. The arm $-Z$ of the repeater adjacent to the line cancels *Z* of the line. The two admittances $-Y$ and *Y* cancel and the other series arms $-Z$ and *Z* also cancel. The result, as shown in Fig. 3(C), is that only the attenuation and phase shift of network *B* remain.

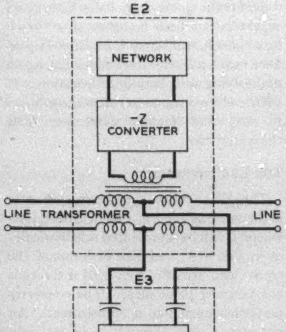

E2
NETWORK
-Z CONVERTER
LINE TRANSFORMER
LINE
E3
-Y CONVERTER
NETWORK

Fig. 5. The bridged T repeater E23

In practice, the amount of attenuation α_1 which can be canceled by the repeater depends on the uniformity, the loss, and the type of line. The permissible magnitude is computed by conventional methods which are beyond the scope of this paper.

Fig. 3 has shown how the combination of a series and a shunt repeater can annul a large part of the attenuation of a telephone line. Much may be accomplished also by a series negative impedance alone as illustrated in Fig. 4. In Fig. 4(A) a transmission line is again represented by two networks *A* and *B*. However, network *A* now is shown as an *L* configuration having a series arm *Z* and a shunt admittance *Y* together with an ideal transformer of ratio 1:*K* where *K* exceeds unity. Network *A* of Fig. 4 is equivalent to network *A* of Fig. 3, and network *B* of Fig. 4 is the same as network *B* of Fig. 3. In Fig. 4(B) the addition to this transmission line of a single $-Z$ such as the *E2* repeater is shown. This negative impedance $-Z$ cancels the series impedance *Z* of network *A*. The result is shown in Fig. 4(C).

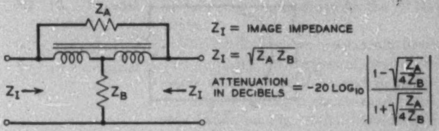

Z_A
Z_I
Z_B
Z_I = IMAGE IMPEDANCE
$Z_I = \sqrt{Z_A Z_B}$
ATTENUATION IN DECIBELS $= -20 \log_{10} \left| \frac{1-\sqrt{\frac{Z_A}{4Z_B}}}{1+\sqrt{\frac{Z_A}{4Z_B}}} \right|$

Fig. 6. Schematic of the bridged T network

Thus Fig. 4 shows that when a single $-Z$ is added in series with the conductors of a transmission line the attenuation is reduced, but the equivalent of a shunt conductance *Y* together with an impedance transformation is left. The impedance transformation 1:*K* could be corrected by means of a transformer if it were not for the shunt element *Y*. Thus an irregularity is introduced which limits the use of a single series repeater.

The Bridged T Structure

The discussion of the *T* repeater, illustrated in Fig. 3, was based on the use of 2-series negative impedances and a shunt negative admittance. It is perfectly possible, and more economical, to obtain the same effect by using a single-series impedance in a bridged *T* structure and this is, in fact, what is done with the *E2–E3* combination repeater. Fig. 5 shows how this is accomplished by using a center-tapped line coil for the *E2* repeater with the *E3* connected as a shunt element to the mid-points of the coil. If the coil is considered to be ideal this arrangement is equivalent to the configuration shown in Fig. 6, in which the coil provides the basic bridge structure: The *E2* repeater is the series arm, and the *E3* the shunt arm. Incidentally, this arrangement of *E2* and *E3* repeaters is similar to Crisson's twin 21-type repeater.[4]

For such a bridged *T* structure, the image impedance equals the square root of the product of the series and shunt

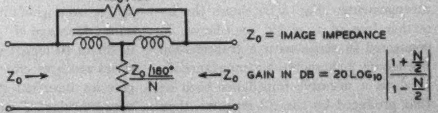

$NZ_0 \underline{/180°}$
Z_0
$\frac{Z_0 \underline{/180°}}{N}$
Z_0 = IMAGE IMPEDANCE
GAIN IN DB $= 20 \log_{10} \left| \frac{1+\frac{N}{2}}{1-\frac{N}{2}} \right|$

Fig. 7 (right). Schematic of the bridged T repeater

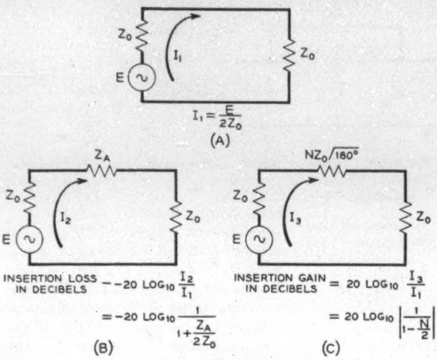

Z₀
I₁
E
Z₀
$I_1 = \frac{E}{2Z_0}$
(A)
Z_A
Z_0
I_2
E
Z_0
INSERTION LOSS IN DECIBELS $= -20 \log_{10} \frac{I_2}{I_1}$
$= -20 \log_{10} \frac{1}{1+\frac{Z_A}{2Z_0}}$
(B)
$NZ_0/\overline{180°}$
Z_0
I_3
E
Z_0
INSERTION GAIN IN DECIBELS $= 20 \log_{10} \frac{I_3}{I_1}$
$= 20 \log_{10} \left|\frac{1}{1-\frac{N}{2}}\right|$
(C)

Fig. 8. Insertion gain of the E2 repeater

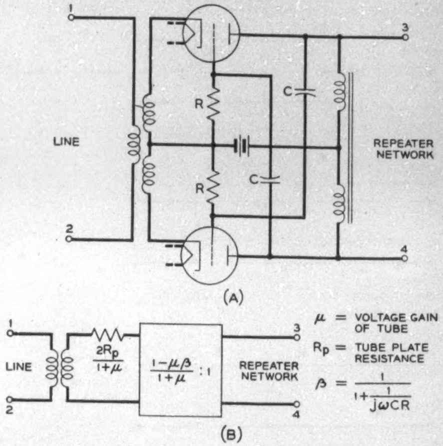

1
2
3
4
LINE
R
R
C
C
REPEATER NETWORK
(A)
1
2
3
4
LINE
$\frac{2R_p}{1+\mu}$
$\frac{1-\mu\beta}{1+\mu} : 1$
REPEATER NETWORK
(B)
μ = VOLTAGE GAIN OF TUBE
R_p = TUBE PLATE RESISTANCE
$\beta = \frac{1}{1+\frac{1}{j\omega CR}}$

Fig. 9. E2 converter

A—Schematic
B—Equivalent circuit

arms and the attenuation, in decibels (db) is as indicated on the drawing.

If a network is to be inserted in an electrically long line without introducing an irregularity, its image impedance must match the characteristic impedance of the line. This would be the case for the bridged *T* network if Z_A were set equal to NZ_0 and Z_B were set equal to Z_0 divided by *N*. Then the square root of the product of Z_A and Z_B would be Z_0.

A network make up of negative impedances is designed to match a line in the same way, and Fig. 7 is a representation of such a structure. Here, as a matter of convenience, the shunt arm is shown as an impedance with a +180-degree phase shift instead of an admittance, as used previously. The letter *N* designates a numeric or proportionality constant. It will be observed that the product of the shunt and series impedances is a real or positive impedance and hence the image impedance is a positive impedance, Z_0. The gain is determined entirely by the value of *N*. Thus if the characteristic impedance of a transmission line is known, together with the gain that the line can support without risk of oscillation, then *N* is known and the repeater network can be adjusted for values of Z_A and Z_B to give the required gain.

The advantage of the bridged *T*, as compared to a single-series negative impedance such as the *E2*, can be demonstrated by comparing the relative transmission gains obtainable from the two arrangements. Fig. 8(B) shows the insertion loss of a single impedance Z_A connected in series with a transmission line having a characteristic impedance Z_0. If Z_A is a negative impedance such as that produced by the *E2* repeater, then the repeater gain becomes a function of *N* as shown in Fig. 8(C). If *N* equals 2 the gain is infinite and the system will oscillate. Thus *N* must always be less than 2 where Z_A is a negative impedance of the series or open-circuit stable type. Practically, the impedance of the transmission line is not a constant Z_0 but varies with termination, line construction, and temperature. Thus *N* should be decreased until the negative impedance is always less than the sum of the two line impedances in series with it, taking into account all possible variations in these impedances.

The same limitations on *N* apply to the bridged *T* repeater of Fig. 7 as apply to the single-series repeater. If *N* is 2 the gain is infinite and the circuit will sing. This similarity goes further. Assume that a single negative impedance Z_A equal to $NZ_0/\overline{180}$ degrees is inserted in series in an electrically long line and *N* is adjusted for stability. If this series element is removed and the bridged *T* of Fig. 7 is inserted in the same place and adjusted by changing *N* until the system is stable, it will be found that *N* will have the same value in the bridged *T* structure as it had for the single-series negative impedance.

Thus, if *N* is the same in the case of the bridged *T* as in the single series impedance, the gain advantage can be obtained by comparing the formulas of Figs. 7 and 8, from which it can be seen that the gain advantage of the bridged *T* is equal in db to $20 \log_{10}[1+(N/2)]$. If a single series repeater can be used in a line to give an insertion gain of 6 db ($N=1$) then a bridged *T* can be used to provide $20 \log_{10}(1+0.5)$ or 3.5 db additional. Thus, in this case the series repeater gives 6 db gain as compared to 6+3.5 or 9.5 db for the bridged *T*. These gains are theoretical; in actual lines with simply constructed repeaters the comparison may not be quite so favorable to the bridged *T*.

The Negative Impedance Converter

So far the discussion of the *E2* and *E3* repeaters has been in terms of a "black box" which translates a positive impedance into a negative impedance through a multiplying and phase shift operation. It will be interesting to examine these boxes to see what factors determine their characteristics.

THE E2 CONVERTER

The *E2* negative-impedance converter is the same as the *E1*. As discussed elsewhere[5] it can be represented schematically as in Fig. 9(A) and also in terms of the equivalent circuit of Fig. 9(B) if the coils are assumed to be ideal. The converter performs much like a transformer. An impedance seen through it is not only transformed in magnitude by the ratio of $|1-\mu\beta|$ to $|1+\mu|$ it is also modified by the phase shift of this factor which, over the operating band of frequencies, approximates 180 degrees. The symbol μ stands for the voltage gain of the vacuum tube and β is the ratio of 1 to $1+(1/j\omega CR)$. If both *C* and *R* are large, β approaches

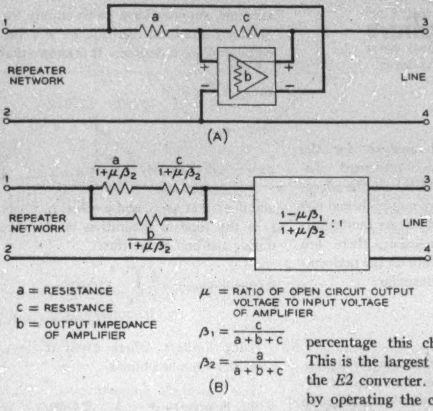

1
a
c
3
REPEATER NETWORK
+
b
+
−
−
LINE
2
4
(A)
$\frac{a}{1+\mu\beta_2}$
$\frac{c}{1+\mu\beta_2}$
1
3
REPEATER NETWORK
$\frac{b}{1+\mu\beta_2}$
$\frac{1-\mu\beta_1}{1+\mu\beta_2}:1$
LINE
2
4
a = RESISTANCE
c = RESISTANCE
b = OUTPUT IMPEDANCE OF AMPLIFIER
μ = RATIO OF OPEN CIRCUIT OUTPUT VOLTAGE TO INPUT VOLTAGE OF AMPLIFIER
$\beta_1 = \frac{c}{a+b+c}$
$\beta_2 = \frac{a}{a+b+c}$
(B)

Fig. 10. E3 converter

A—Schematic
B—Equivalent circuit

unity in magnitude and the ratio of conversion approaches $1-\mu$ to $1+\mu$. If μ is large compared to unity, then this conversion ratio approaches -1. This ratio of -1 is approximately realized in the *E2* converter, and therefore the conversion ratio is not changed appreciably by small variations in μ.

In addition to the transformation term there is also, as shown in Fig. 9(B), a series term, $2R_p$ divided by $1+\mu$. Here R_p is the plate resistance of the tube, and μ is the voltage gain, as mentioned before. The factor 2 results from the use of two tubes in push pull. If μ is large compared to unity then this series term becomes approximately $2R_p/\mu$. It is entirely dependent upon the characteristics of the vacuum tube. As the characteristics change from tube to tube with manufacturing variation or in the same tube over a period of time or with variation in battery supply potential, the term $2R_p(1+\mu)$ will change accordingly. In percentage this change may be large. This is the largest source of variation in the *E2* converter. It can be minimized by operating the converter between impedances much larger in magnitude than $2R_p/(1+\mu)$ so that variations in this term have relatively small effect. This has been done in the *E2* repeater by stepping up the impedance of the transmission line by about 1:9 by means of the transformer shown in Fig. 9.

THE E3 CONVERTER

Theoretically, the same converter used for the *E2*, and shown in Fig. 9, could have been used for the *E3*. Instead of connecting the line to the converter through terminals 1 and 2, terminals 3 and 4 would be used. However, because the *E3* must be designed for connection across a transmission line a coil or transformer input is not practical since the coil would shunt the line at low frequencies and introduce excessive loss to dial pulsing and 20 cycles per second ringing. Without a coil to step up the impedance of the line, variations in $2R_p/(1+\mu)$ with standard triodes are too large to be neglected. For this reason another converter circuit was designed for the *E3* repeater.

This circuit is shown in schematic form in Fig. 10(A). It consists of two resistances, *a* and *c* respectively and an amplifier poled according to the plus and minus designation on Fig. 10(A). The output impedance of the amplifier has been designated as *b*. If the input impedance of this amplifier is high compared to other circuit impedances, Fig. 10(A) can also be represented by the equivalent circuit of Fig. 10(B). Here is a conversion factor similar to that in the *E2* converter and also a series impedance. In Fig. 10(B) the factor μ is the ratio of the open-circuit output voltage to the input voltage of the amplifier. In the *E3* converter this voltage ratio μ is quite high because the amplifier is a 2-stage arrangement. In the design of the *E3* both β_1 and β_2 are approximately one half. Thus $\mu\beta_1$ and $\mu\beta_2$ are both large compared to unity so that the conversion ratio $(1-\mu\beta_1)/(1+\mu\beta_2)$ is approximately unity and relatively independent of variations in μ. Furthermore, because $\mu\beta_2$ is large compared to *b* the series term in the converter circuit is relatively small, and variations in this term have little effect on the operation of the converter.

References

1. A NEGATIVE IMPEDANCE REPEATER, **J. L. Merrill, Jr.** *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1461–66.

2. E-TYPE TELEPHONE REPEATERS—DESCRIPTION, EQUIPMENT, AND TESTING, **J. O. Smethurst.** *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 435–43.

3. NEGATIVE IMPEDANCE TELEPHONE REPEATERS—APPLICATION IN THE BELL SYSTEM, **Arthur F. Rose.** *Ibid.*, pp. 430–35.

4. NEGATIVE IMPEDANCE AND THE TWIN 21-TYPE REPEATER, **G. Crisson.** *Bell System Technical Journal*, New York, N. Y., July 1931.

5. THEORY OF THE NEGATIVE IMPEDANCE CONVERTER, **J. L. Merrill, Jr.** *Bell System Technical Journal*, New York. N. Y., Jan. 1951.

No Discussion

Magnetic Frequency Multipliers

L. J. JOHNSON
ASSOCIATE MEMBER AIEE

S. E. RAUCH
NONMEMBER AIEE

WITH the availability of high-quality magnetic core materials, it has become possible to develop very efficient magnetic frequency multipliers. The advance of the electronic arts in the development of computers and precise control systems has created the immediate need for frequency multipliers which can produce exact multiplication of frequencies in power levels ranging from milliwatts to one kilowatt. Magnetic frequency multipliers as described in this paper have superior characteristics to electronic multipliers in that they use no tubes, moving parts, or rectifiers. Consequently, they are rugged, efficient, insensitive to temperature variations, and can deliver large quantities of power.

The use of the saturable core reactor in frequency multiplication circuits was of considerable interest in the earlier days of the development of radio-frequency transmitters. The core materials available at that time represented a low quality magnetic steel, and consequently the designs developed yielded units which were not only inefficient but large in weight and volume. In spite of the poor materials it was possible to obtain frequency multiplication of very high magnitudes. Typical designs and applications of that period are to be found in references 1–5.

The saturable reactor core has made rapid changes in its characteristics as the advances in high-quality rectangular-loop magnetic materials have been placed on the market. As a result, the application of magnetic amplifier circuits has grown extensively. An equally interesting as well as promising advance for the saturable reactors can be found in the field of static-frequency multiplication. The fact that the rectangular loop core materials yield a significant percentage of high harmonics when operating across the knee of the saturation region suggests that very efficient frequency multipliers of high values of multiplication can be developed.

The mathematical analysis for the general design theory is presented. Included in the design theory are the effects of various types of core materials and circuitry for selected odd-integer multiplication ratios. In conclusion, there are shown experimental data on the performance of a standard quintuplar.

Design Theory

For the better understanding of the discussion that follows on the design of frequency multiplier circuits, let us examine first the simple series circuit employing a single saturable reactor as shown in Fig. 1. Fig. 2 represents the corresponding supply and load voltages versus time t in radians when the a-c supply voltage is adjusted to a value which will saturate the core. It is well known that the behavior of the load voltage is dependent upon the magnitude of the supply voltage and the gating effect of the saturable reactor.

Let $E(t)$=applied voltage; $\Phi(t)$= instantaneous total flux in the core; t_0= time at the beginning of a half cycle, and the steady state operation has been reached. Thus at the beginning of a half-cycle, the voltage $E(t)=0$ and is going positive. The flux $\phi(t)$ in the core is at its negative remanance value. As time progresses the supply voltage $E(t)$ increases in a sinusoidal manner and the flux $\phi(t)$ of the core increases to positive saturation at the approximate rate $d\phi/dt=E(t)/N\times 10^8$.

At saturation the rate of change of flux rapidly approaches zero. The voltage across the saturable reactor likewise approaches zero, and the load voltage approaches the supply voltage $E(t)$ for the remainder of the half-cycle. This condition is obtained only if the volt-second magnitude of the supply voltage is greater than the flux linkage magnitude of the saturable reactor. Thus one obtains the characteristic as shown in Fig. 2. The time within the half-cycle at which the core saturates is called the firing angle, θ_f, and is conveniently defined in terms of degrees or radians. θ_f is completely determined by the volt-seconds of the half-cycle and the flux linkages at saturation of the core.

The firing angle can be adjusted to any preassigned value by the proper selection of volt-seconds and flux linkages. For example, suppose for a given supply frequency f and total flux linkage Φ a specific θ_f radians is desired. It is known that

$$E(t)=N\frac{d\phi}{dt}\times 10^{-8}$$

or

$$d\phi=\frac{1}{N}\times E(t)dt\times 10^{8}$$

and if $\phi=0$ at $t_0=0$, and $\phi=\Phi$ at t_f, where t_f is the time of saturation of the core during the half-cycle, then

$$\Phi\ (\text{flux linkages})=\frac{1}{N}\int_{t_0}^{t_f} E(t)dt\times 10^{8}\ (\text{volt-seconds})$$

Using $\theta=2\pi ft$, where $\theta=0$ at $t=t_0$, $\theta=\theta_f$ at $t=t_f$, one obtains

$$\Phi\ (\text{flux linkages})=\frac{1}{2\pi fN}\int_{0}^{\theta_f} E(\theta)d\theta\times 10^{8}\ (\text{volt-seconds})$$

For the important application of sinusoidal voltage $E(t)=E_{\max}\sin 2\pi ft$

$$\Phi=\frac{E_{\max}}{2\pi fN}\int_{0}^{\theta_f}\sin\theta\, d\theta\times 10^{8}$$

$$\Phi=\frac{E_{\max}}{2\pi fN}[1-\cos\theta_f]\times 10^{8}$$

and

$$E_{\max}\ (\text{volts})=\frac{2\pi fN}{1-\cos\theta_f}\phi\times 10^{-8}$$

It is noted that the initial time $t=0$ was determined by the supply voltage $E(t)=E_{\max}\sin(2\pi ft)$. To obtain multiplication it is necessary to use separate supply voltages which are equal in magnitude but shifted in phase by multiples of the angle θ_f. For all odd numbers of multiplication the supplies can have a common ground and a single load. An example of such a circuit is shown for a tripler in Fig. 3. It can be seen that no two reactors can be saturated at the same time, for if they were the supplies would be short-circuited. For this reason, it is required that $n(\pi-\theta_f)=\pi$. In the general case for frequency multiplication by an odd integer n, for which the supplies have

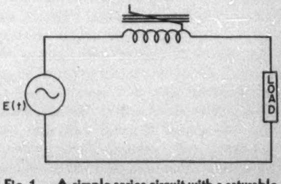


Fig. 1. A simple series circuit with a saturable reactor

Paper 54-248, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 22, 1954; made available for printing May 4, 1954.

L. J. JOHNSON is with the Hufford Machine Works, Los Angeles, Calif.; and S. E. RAUCH, Professor of Mathematics, is with the University of California, Santa Barbara, Calif.

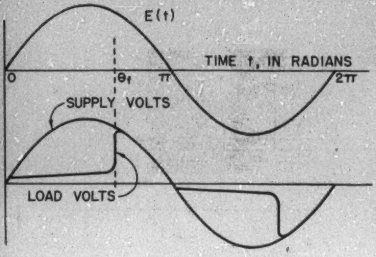


Fig. 2 (left). Supply and load voltages versus time for the circuit of Fig. 1

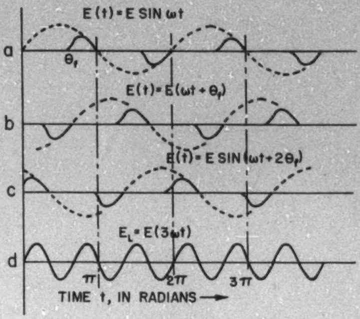


Fig. 4 (right). Composite wave forms which yield frequency tripling for the circuit in Fig. 3

a common ground and single load, the firing angle is determined by

$$\theta_f = \pi\left(\frac{n-1}{n}\right)$$

and the n distinct supplies are phased such that

$$E_1 = E_{\max} \sin (2\pi ft)$$
$$E_2 = E_{\max} \sin (2\pi ft + \theta_f)$$
$$E_3 = E_{\max} \sin (2\pi ft + 2\theta_f)$$
$$\vdots$$
$$E_n = E_{\max} \sin (2\pi ft + [n-1]\theta_f)$$

A graphical illustration showing the manner in which frequency tripling multiplication results is provided in Fig. 4. The corresponding circuit in Fig. 3 shows three separate supplies shifted in phase $2\pi/3$ radians. Thus the firing angle for any of the three phases is fixed at $2\pi/3$ radians, for optimum conditions. The curves a, b, and c in Fig. 4 represent the supply voltages in dashed lines and the contribution to load voltages of each supply in solid lines. The voltage across the load is the sum of the contributions of the separate phases, and the resultant is shown as c in Fig. 4. It should be noted that each reactor fires only once during the half-cycle of its own supply.

Frequency triplers are easily made since 3-phase supply systems are commonly used. Phase-changing transformers can be used with 2- or 3-phase supplies to obtain higher odd multiples of phases required for higher frequency multiplication values.

Even integer frequency multiplication is obtained by using circuits containing rectifiers, d-c bias circuits, or permanent magnet bias. For even multiples the weight and volume of multiplier units are larger for a given power rating than the corresponding values for odd-integer multiplier units. This is a result of the use of biased cores which are saturated at only one end of the hysteresis loop. Therefore, the number of reactors is doubled for each phase. Even integer multiplication will be considered in detail in a future paper.

Discussion

Wave Form

Core Materials

The output voltage wave form of the frequency multiplier is dependent upon the saturation characteristic around the knee of the hysteresis loop of the core material employed. Fig. 5(B) is a representative plot of hysteresis loop sections of three well-known magnetic core material types. The rate at which saturation occurs is markedly different for the three cases shown.

In Fig. 5(A) there are shown the supply voltage wave form $E(t)$ and the normal flux wave form $\phi(t)$ as basic parameters. Superimposed on these two curves are the load voltage characteristics and the flux time characteristics for the three core materials in Fig. 5(B). The wave form of the load voltage when using core A is a chopped sine wave. Core B yields a load voltage wave form with reduced high frequency harmonics, and core C provides a load voltage wave form which is most nearly sinusoidal. Fig. 5(C) is a current-time relationship of load current $I(t)$ for the three core materials.

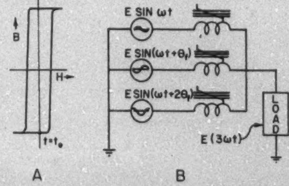


Fig. 3. Three separate supplies shifted in phase $2\pi/3$ radians

A—Rectangular hysteresis loop core material characteristic
B—Schematic diagram of a tripler

Although Fig. 5(A) indicates a most nearly sinusoidal wave form for core C, the reader should not be misled by this conclusion, since θ_f changes as the multiplication factor n increases. For a tripler, i.e., $n=3$, core C would yield a near optimum wave shape. For the quintuplar, core B is more appropriate since the firing angle θ_f is closer to the value π. Thus, one must reproduce a sine curve in a smaller portion of the supply wave which will require a more peaked saturation spike. In general, as the ratio of load frequency to supply frequency increases, the knee of the saturation curve must become more rectangular.

Loading Effects

Although the wave form is not appreciably affected by linear loads, the nonlinear loads such as magnetic amplifiers and biased rectifiers will alter the wave form on an impedance ratio basis. Therefore, it is desirable to maintain the multiplier internal impedance low on a dynamic basis. This can be accomplished in design by the use of low internal winding resistance, which increases the weight and volume of the unit. Another method available is the use of tuned output circuits.

Removal of Undesired Harmonics

Although undesired harmonics in the load currents which are the results of nonlinear loads can be reduced by the two methods described in the preceding section, only the tuned output circuit method will reduce undesired harmonics which are generated within the multiplier. Tuned output circuits can be applied successfully with output wave forms which have maximum total harmonic distortion of 0.5 per cent.

The selection of appropriate core ma-

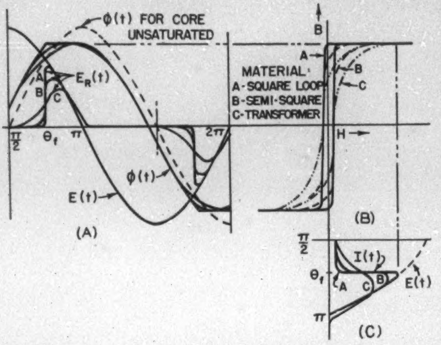


Fig. 5. Effects of various core materials on output voltage wave form

A—Plot of voltages and flux versus time in radians
B—Typical hysteresis loops for 3-core materials
C—Plot of load currents versus time in radians

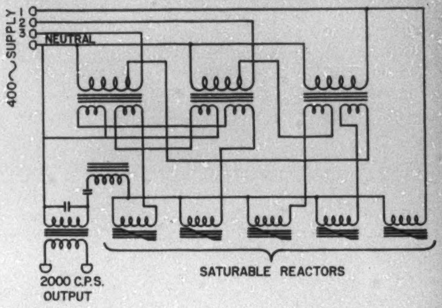


Fig. 6 (above). Circuit schematic for a frequency quintuplar

terials and low impedance windings can yield designs which produce output wave forms with less than 5 per cent total harmonic distortion, without the use of tuned output circuits.

Regulation

The regulation of the load voltage and frequency is related closely to the regulation of the supply. The dynamic characteristics of the output can be improved considerably for load transients by the use of tuned output circuits. However, the additional d-c resistance contributed by special tuning devices is detrimental to optimum steady-state regulation. Steady-state regulation is held to a minimum by the careful selection of design parameters, including winding resistance, core materials, and reduction of eddy currents.

Since frequency multipliers operate on the saturation principle, it is necessary to limit the regulation of supply voltage and frequency. The optimum operation demands close regulation on both supply voltage and frequency. The overvoltage tolerance of frequency multipliers is determined mainly by the multiplication factor, the desired operating efficiency, and the output wave form tolerance. As has been indicated previously, the saturation characteristic of the optimum core material for a given multiplier is determined by its multiplication constant *n*. Furthermore, if at any time two cores are saturated, the corresponding supply phase will be short-circuited, and the short-circuit currents not associated with the load can flow through the two reactors. Since frequency multipliers with low values of *n* normally employ core materials of the type shown in Fig. 5(B), curve *C*, it follows that supply voltage regulation tolerance is at a maximum; consequently, the rectangular core characteristics require the closest supply voltage regulation. Overcurrent protection should be provided in each phase of the supply voltage as well as in the load circuit to obtain maximum multiplier protection.

Efficiency

The efficiency of frequency multipliers, as in all other power devices is determined by the internal losses and the delivered power. A well-designed frequency multiplier will have in general power efficiencies in the range of 50 to 90 per cent over its useful load range. Wave form correction devices such as tuned output circuits decrease power efficiency, since unwanted harmonics must be dissipated within the multiplier. For this reason it is desirable to design multipliers which do not require tuned output devices. Since the internal multiplier losses are primarily resistive, the best design procedures should minimize the winding resistances.

Table I. Quintuplar Specifications

Input	
Supply Voltage	125 volts, 3-phase
Supply Frequency	400 cycles per second
Supply Regulation	±2 per cent on voltage and frequency
Output	
Voltage	125 volts, single-phase
Frequency	2,000 cycles per second
Power	3 watts
Duty Cycle	continuous
Distortion	less than 1 per cent total harmonic distortion
Impedance	low as practicable

In many applications the reduction of weight and volume is an important consideration. Methods which will reduce the weight and volume include single windings on saturable reactors, careful choice of core materials, and the use of reasonable tolerances for harmonic distortion. Single coils on reactors allow the entire window to be used for one winding, thereby optimizing the power rating of a given reactor. A carefully chosen core characteristic will maximize the desired harmonic content at the desired frequency which minimizes the over-all core volume. Increased unwanted harmonic tolerances can eliminate the need of external tuning components.

Experimental Evaluation

The authors have designed and built frequency multipliers with power ratings ranging from 1 to 300 watts and with multiplication values n ranging from 3 to 10. The application of the principles described in the foregoing design theory was used to obtain high efficiency and good wave form. Multiplier efficiencies having magnitudes as high as 90 per cent were achieved. An increase in weight and volume over rotary equipment which yielded equivalent power and frequency ratings was noted in all units, but the advantageous characteristics of reliability, long life, and almost no maintenance problems were evident for the magnetic frequency multipliers. A specific example of frequency multiplication is the quintuplar having the specifications shown in Table I.

The circuit for this multiplier is shown in Fig. 6. It is necessary to incorporate transformers in the 400-cycle supply cir-

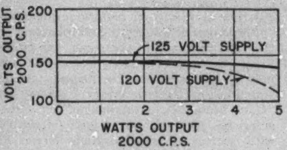


Fig. 7. Output characteristics for a 3-watt quintuplar

cuit in order to change the three phases to five phases. The saturable reactor circuit is similar to that shown in Fig. 3(B), except for the use of the 5-phase star circuit. Since the firing angle θ_f should be $4\pi/5$ radians, a core material having a moderately rectangular hysteresis loop is desirable. Fig. 5(B), curve *B*, is a representative plot of the desirable core material characteristic.

The requirement for the total distortion not to exceed 1 per cent necessitates the introduction of a tuned output circuit since the core material chosen produces distortions in the order of 5 to 10 per cent. It may be noted further that the output circuit includes a transformer which is used to obtain the rated output voltage of 125 volts at rated load. The resonant transformer as used provides both the proper voltage as well as low impedance to transient loads. Although the transient impedances are low, the impedance for steady-state conditions is relatively unaffected by the tuned output circuit.

The losses in the basic multiplier unit are the sum of the winding losses and core material losses. The circuit as shown in Fig. 6 has a particular advantage in reducing internal losses since only one winding per reactor is used. The single winding can occupy completely the core window, thereby allowing a larger wire size for the reduction of resistance. Previous multipliers, such as described in reference 6, have employed wye-to-delta transformation requiring two windings per core which inherently increases copper losses by a factor of four.

Tests on the unit just described yielded a load characteristic curve as shown in Fig. 7. Although the curves for Fig. 7 describe the behavior of a specific quintuplar, they can be considered to be typical for frequency multipliers of this type. The output wave form for the 3-watt quintuplar had a measured toal harmonic distortion of 0.5 per cent. The over-all dimensions, including transformers and resonant circuits, were approximately 3 by 3 by 5 inches, and the weight was approximately 4 pounds. The volume and weight of frequency multipliers do not increase linearly with power rating. For example, a 25-watt quintuplar, 400–2,000 cps, employs the same volume and has approximately the same weight. A 300-watt quintuplar for the same frequency requirements will need 4 by 6 by 7 inches and will have a weight of approximately 30 pounds.

Conclusions

When design procedures as set forth in this paper are observed carefully, magnetic frequency multipliers can be built which have reliability and long life equivalent to power transformers. The efficiencies can be maintained within a 50 to 90 per cent range for high multiplication values. Although commercially available units are larger and heavier than their equivalent rotary high-frequency generators, they are to be preferred where space and weight are not the prime requirements.

The effect of a low power factor load on frequency multipliers is a shift in the firing angle of reactors referred to their respective supplies. Since the reactor of each phase has a fixed volt-second magnitude, the presence of a reactive component in the load will not affect the volt-ampere output rating of the multiplier.

The details of this paper have dealt with odd-integer frequency multiplication. It is not to be inferred that magnetic frequency multiplication is limited to the odd-integer multiplications. The design details for even-integer multipliers are similar in nature but require different circuit techniques to obtain optimum performances. Design details for even-integer multipliers will be presented in the future.

References

1. Frequency Multiplication—Principles and Practical Applications of Ferro-Magnetic Methods, **N. Lindenblad, W. W. Brown.** *AIEE Transactions*, vol. 44, Apr. 1925, pp. 491–496.

2. Radio Frequency Changers, **A. N. Goldsmith.** *Proceedings*, Institute of Radio Engineers, New York, N. Y vol. 3, March 1915, pp. 55–79.

3. A Contribution to the Theory of Magnetic Frequency Changers, **J. Zennech.** *Ibid.*, Dec. 1920, pp. 468–92.

4. Static Frequency Multipliers for the Production of Very High Frequencies in Radio Telegraphy, **M. Latour.** *Revue générale de l'électricité*, Paris, France, July 1922.

5. Contribution to Frequency Transformation by Means of Iron Core Inductance, **Dornig.** *Elektrotechnische Zeitschrift*, Wuppertal-Elberfeld, Germany, vol. 45, no. 42, Oct. 1924, pp. 1107–08.

6. Magnetic Frequency Conversion, **L. C. Harriott.** *Proceedings*, National Electronic Conference, Chicago, Ill., vol. 9, 1953, pp. 78–87.

◆

Discussion

L. A. Finzi and **G. C. Feth** (Carnegie Institute of Technology, Pittsburgh, Pa.): At the start of the paper it is recalled that magnetic frequency multiplication circuits have been known for many years. It may be of historical interest to mention that the earliest disclosure of a frequency multiplier of the type shown in Fig. 3(B) is found in a patent issued in 1912.[1] This does not detract from the merit of the authors, who have most judiciously chosen a proper range of frequencies and powers for their multipliers and have obtained remarkably practical results.

We do not quite agree, however, with the design theory given. It seems to us that the choice of an angle of firing $\theta_f = [(n-1)/n]\pi$ in a frequency multiplication by n does not prevent intervals of simultaneous saturation of reactors in pairs in the steady state. Under the assumption of low winding resistances, of stiff voltage sources, of resistive loads and of core materials sensitive enough to allow for an analysis in terms of "angles of firing," it seems rather necessary to use values of θ_f larger than those indicated, if short circuits of the supplies are to be avoided.

Consider for instance Fig. 3(B) with symmetrical line-to-neutral source voltages $e_a = E_{\max} \sin \omega t$, $e_b = E_{\max} \sin (\omega t - 120$ degrees) and $e_c = E_{\max} \sin (\omega t - 240$ degrees). (These are the voltages represented by the dashed lines of Fig. 4.) Shortly before ωt reaches 180 degrees reactor *a* is in saturated conditions while the other two cores, unsaturated, are absorbing voltages. At $\omega t = 180$ degrees the voltage e_a goes to zero and then reverses, but this is not sufficient to remove reactor *a* from saturation if, by the assumed choice $\theta_f = 120$ degrees, reactor *c* fires just at this time. With *c* saturated a short circuit current flows in the loop of the reactor windings *a* and *c* and the direction of this current is such to prevent reactor *a* from entering its unsaturated region. This short circuit persists until $|e_a| = |e_c|$, i.e., until $\omega t = 210$ degrees; it is only from this point that e_a prevails upon e_c and the reactor *a* starts absorbing (negative) volt time areas.

To allow reactor *a* to stop conducting at $\omega t = 180$ degrees it is necessary to delay the firing of core *c*. Then a process of negative volt time areas absorption may well initiate in core *a*; subsequently, if *c* fires at any time for which $|e_c| > |e_a|$ this process may well reverse, that is, core *a* is pushed back toward the initial positive saturation and may or may not re-fire. It can be shown that if $\theta_f = 137$ degrees, core *a* barely reapproaches positive saturation at $\omega t = 210$ degrees and from then on it is steadily brought down toward negative saturation, under the combined action of e_a and of the voltage which appears across the load because of saturation of *b* or *c*. It is found that the crest of phase voltage $E_{\max}$ (volts) is related to the turns *N* of reactor winding and to the core saturation flux ϕ_{sat} (webers) by $E_{max} = 1/\sqrt{3}\,\omega N 2\phi_{sat}$ for $\theta_f = 137$ degrees.

We are discussing these overlap phenomena and their prevention at length because the matter of simultaneous saturation of reactors is of concern not only in the design of frequency multipliers but also in the broader field of polyphase magnetic amplifier applications. (In fact, similar problems occur also in polyphase switching circuits, thyratrons etc.)

The authors have recognized the usefulness of less sensitive cores as a means to repress overlap short circuits. But the analysis and design theory becomes much more involved when clean-cut concepts of firing are abandoned. This makes the practical information given in the paper even more interesting.

REFERENCE

1. **F. Spinelli.** Italian Patent Office Patent No. *124825*, 1912.

L. J. Johnson and **S. E. Rauch**. The authors wish to thank the discussers for their worthwhile addition to the paper. For the theoretical case in which remanence flux is equal to the saturation flux, they are correct in their analysis of the short-circuit conditions which would exist for a tripler between the angles of 120 to 137 degrees.

Rectangular core materials which are commercially available today have a remanence flux ranging from 5 to 10 per cent less than saturation flux measured at one oersted, as quoted by Arnold Engineering Company for their Deltamax material. The average remanence flux is approximately 7 per cent below the saturation flux measured at one oersted. In the practical application of the best rectangular core materials to magnetic frequency multiplication, the flux change above remanence is sufficient to limit the cross-fire currents to magnitudes considerably less than rated load currents. For this reason the firing angle defined as $\theta_f = [(n-1)/n]\pi$ is a realistic evaluation.

For magnetic frequency multiplication of the type discussed in the paper, the cross-firing effects due to sharp saturation are most pronounced for the tripler. As is stated in the third paragraph of the section "Core Materials," rectangular flux characteristics become increasingly desirable with increasing multiplication factor n. As a consequence, the allowable limit of flux change above remanence can be reduced as the multiplication factor increases.

The authors agree with Mr. Finzi and Mr. Feth that cross-firing is very important in multiphase magnetic amplifiers. The firing angle is controlled and varied in such cases by resetting the core while rectifiers block the current flow in the load circuit. Cross-firing limits the available reset time; therefore it reduces the effectiveness of the applied control voltage. In contrast, the magnetic frequency multiplier has no similar control reset function.

Accurate Tachometry Methods with Electronic Counters

J. M. SHULMAN
ASSOCIATE MEMBER AIEE

Synopsis: High-speed electronic counters have become commercially available within the past few years which indicate by means of a display of neon-lighted figures, and can count from 20 to 100,000 events per second or more with an inherent accuracy of ±1 count in the measuring interval. By relatively simple modifications of the basic counting circuits in these instruments they can also be used for measuring short intervals of time to an accuracy of ±10 microseconds. The measured time interval is displayed directly in figures indicating the decimal fraction of a second to the nearest hundred-thousandth. Tachometry is a fruitful application of these instruments. Three methods of using the counters in tachometry are described here, and the advantages and limitations of each method are discussed.

A PERSISTENT problem of making rapid, accurate speed measurements on a large motor test floor led to the investigation of the commercially available electronic counter as a tachometer. In developing a form of tachometer suitable for the particular test floor conditions involved, three tachometry methods were investigated. Since each of the three systems has advantages and disadvantages for any given application, all three will be discussed with the aim of pointing these out and indicating thereby the particular applications in which each might be most useful. The importance of this problem is attested by the considerable literature on electronic tachometry.[1–6] However, prior to the use of the decimal electronic counter[7] most of the arrangements were too complicated and cumbersome for general use.

Eight criteria were set up as a basis for determining the relative merit of different tachometry methods:

1. Can measurements be obtained at any speed within the required range to the required accuracy?
2. Can measurements be obtained without mechanical coupling to the shaft?
3. Is human error minimized?
4. Can measurements be taken by one person instead of two?
5. Can the equipment be used by relatively unskilled personnel?
6. Can the equipment, particularly the pickup device, withstand mechanical abuse?
7. Can a change in measurements from one shaft to another be made easily and quickly?
8. Is the reading obtained directly in revolutions per minute (rpm)?

All the tachometry methods used prior to the electronic counter failed to meet one or more of these requirements. The first two electronic counter methods to be described, while not meeting all the requirements, filled some of them outstandingly well. The third method is believed to be an approach to satisfying all of them.

Tachometry by Time Interval Measurement

In this method of determining shaft speed, the time interval for 1 revolution of the shaft is measured and displayed in hundred-thousandths of a second by the counter. Fig. 1 shows the block diagram of the system for doing this. The output of a 100-kc oscillator is fed through an electronic gate to the counter input. A magnetic pickup placed near the keyway on the shaft produces a signal which starts and stops the electronic gate on successive signal pulses produced as the keyway passes the pickup. The number of cycles of the 100-kc oscillator output permitted to pass through to the counter within the period of 1 revolution T_R is thus numerically equal to T_R in hundred-thousandths of a second. This number is displayed by the counter and repeated at any preset interval. The shaft speed in revolutions per second is the reciprocal of the counter reading times 10.[5]

The time interval method has an inherent accuracy inversely proportional to the measured speed, as shown in Fig. 2.

Paper 54-290, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted August 21, 1953; made available for printing April 27, 1954.

J. M. SHULMAN is with the Westinghouse Electric Corporation, Sunnyvale, Calif.

Acknowledgment is made to Dr. W. A. Edson of the Stanford University Electronics Research Laboratory for helpful suggestions regarding multiplier circuitry and for assistance in editing the text.

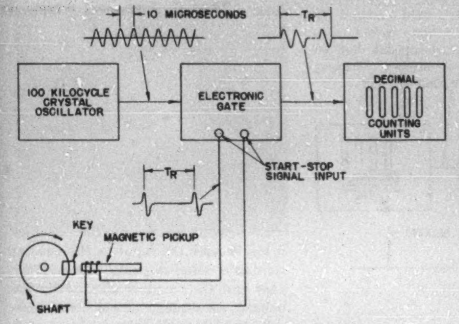

10 MICROSECONDS
T_R
100 KILOCYCLE CRYSTAL OSCILLATOR
ELECTRONIC GATE
DECIMAL
COUNTING UNITS
START-STOP SIGNAL INPUT
T_R
KEY
MAGNETIC PICKUP
SHAFT

Fig 1 (above). Block diagram of system for determining shaft speed by measurement of time interval for 1 revolution

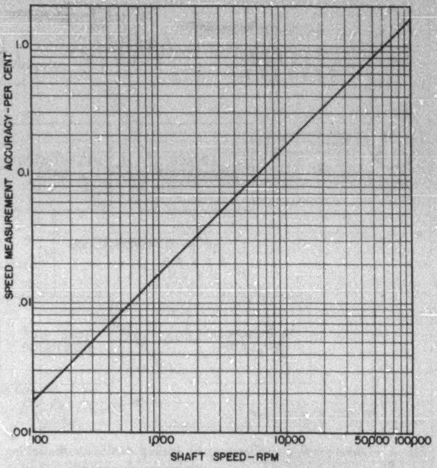


Fig. 2 (right. Theoretical maximum accuracy of speed measurement by time interval

This is the theoretical accuracy based on a counting accuracy of ±1 count in the interval T_R. Other errors in the system may raise the line slightly, but over the speed ranges most commonly measured, between 100 and 10,000 rpm, the order of accuracy obtainable is higher than that possible with most nonlaboratory tachometry methods. At very high speeds it would be possible to increase the accuracy by use of a higher frequency oscillator or by counting for more than 1 revolution. The high accuracy of this method, particularly at speeds of 3,600 and lower, is its primary advantage. Its disadvantage is that it is not a direct-reading method.

Direct-Reading Tachometer with Mechanical Multiplier

Fig. 3 shows the block diagram of an arrangement mechanically obtaining and then counting 60 times the number of shaft revolutions per second and thus displaying on the counter the exact rpm of the shaft. A 60-tooth steel disk mounted on the shaft is used with a magnetic pickup to give a number of output pulses per second numerically equal to rpm. These pass through an electronic gate to the counter. The gate is opened and closed by two successive pulses spaced at exactly 1 second from a time base generator. Depending on the accuracy of the time base desired, the 1-second time base can be derived from either a 100-kc oscillator or from a 60-cycle line frequency.

The main disadvantage of this method is the necessity of having a portion of shaft extension available on which the disk can be attached. Also the problem of moving the disk from one shaft to another may be a disadvantage if motors having many shaft sizes are to be tested. The practical importance of this particular point as it relates to a motor production test floor is worth stressing. In the laboratory, it is considered no trouble at all to put an attachment on a shaft for tachometry purposes. On a test floor, however, the operation is time-consuming, and test personnel will tend to fall back on more convenient methods rather than put on and take off disks even when adequate adapters are available. The desire to eliminate the disk was the main motivating factor which led to the development of the third method.

Direct-Reading Tachometer with Electronic Multiplier

As shown by the block diagram in Fig. 4, this is a method in which the counter

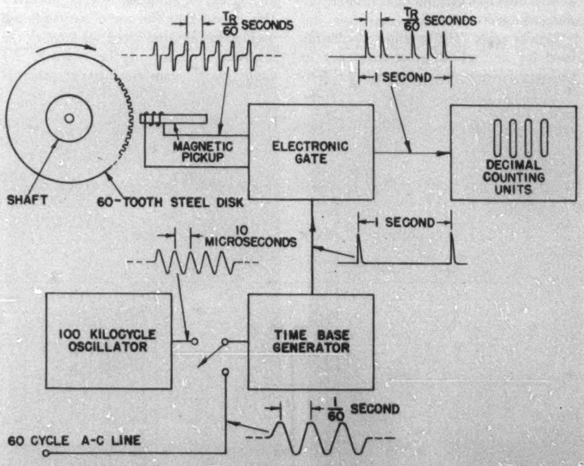

$\frac{T_R}{60}$ SECONDS
$\frac{T_R}{60}$ SECONDS
1 SECOND
MAGNETIC PICKUP
ELECTRONIC GATE
DECIMAL COUNTING UNITS
SHAFT
60-TOOTH STEEL DISK
10 MICROSECONDS
1 SECOND
100 KILOCYCLE OSCILLATOR
TIME BASE GENERATOR
$\frac{1}{60}$ SECOND
60 CYCLE A-C LINE

Fig. 3. Direct-reading tachometer with mechanical multiplication to indicate 60 times the revolutions per second

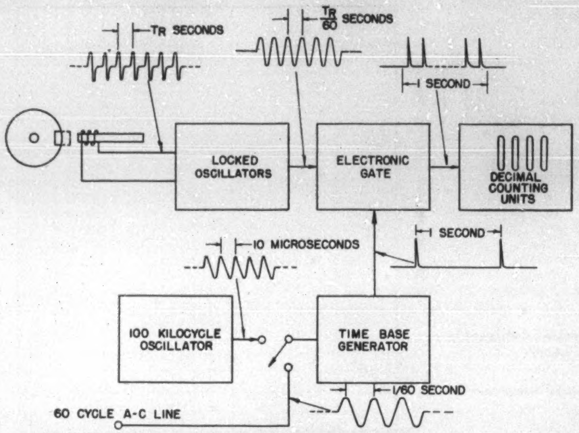

TR SECONDS
$\frac{T_R}{60}$ SECONDS
1 SECOND
LOCKED OSCILLATORS
ELECTRONIC GATE
DECIMAL COUNTING UNITS
1 SECOND
10 MICROSECONDS
100 KILOCYCLE OSCILLATOR
TIME BASE GENERATOR
1/60 SECOND
60 CYCLE A-C LINE

Fig. 4. Direct-reading tachometer with locked oscillators controlling a count of 60 times the revolutions per second

counts and displays the number of cycles per second of the last of a series of sinusoidal oscillators. Each oscillator in the series is locked by one of lower frequency, and the first one in the series is locked by a 1-pulse-per-revolution signal from the shaft. If the total multiplication of the series is 60, the frequency of the last oscillator is numerically equal to the shaft speed in rpm. The arrangement shown in Fig. 4 in effect substitutes an electronic circuit for the 60-tooth disk in Fig. 3, and thus gives a direct-reading method without the disadvantages of the disk.

Synchronized free-running oscillators used for multiplication of frequency in this way can actually accomplish the division of a time interval, in this case the period of 1 revolution of the shaft, into equal parts. This is essentially the same action that the disk accomplishes mechanically when its teeth are equally spaced around its periphery. On the other hand, electronic multiplying circuits which produce a discrete number of output signals for each input signal cannot do this.[8] They would correspond in the mechanical analogy to a disk having its teeth extending over only a portion of its periphery. In terms of speed-measuring accuracy this means that the use of synchronized oscillators can give speed readings to the inherent accuracy of the counter, ±1 count in the counting interval, whereas with straight multipliers the accuracy is less because the input to the counter consists of series of pulses separated by gaps and the count is affected by the point in the sequence at which the measuring period starts.

It is convenient to use locking ratios of 3, 4, and 5 to obtain a total multiplication of 60. Fig. 5 shows a circuit using phase-shift oscillators[9-12] with circuit constants for locking 3,600-cycle nominal output with 60-cycle input. In tests to determine the frequency ranges over which these circuits will lock at different ratios, the results shown in Table I were obtained:

Table I. Frequency Ranges for Different Ratios

Locking Ratio	Locking Range, Per Cent of Highest Frequency
5 to 1	13.5
4 to 1	14.8
3 to 1	22.6
2 to 1	25.2

To cover a speed-measuring range greater than the lowest locking range obtainable, it is necessary to provide means for changing the oscillator frequencies either continuously or in switched steps. The latter method is convenient when speed measurements are to be made on a-c induction motors only, since the slip within the loading range is usually less than 10 per cent.

Signal Pickup

A satisfactory pickup for getting signals from a shaft keyway, or toothed disk, is essential to the success of any of the

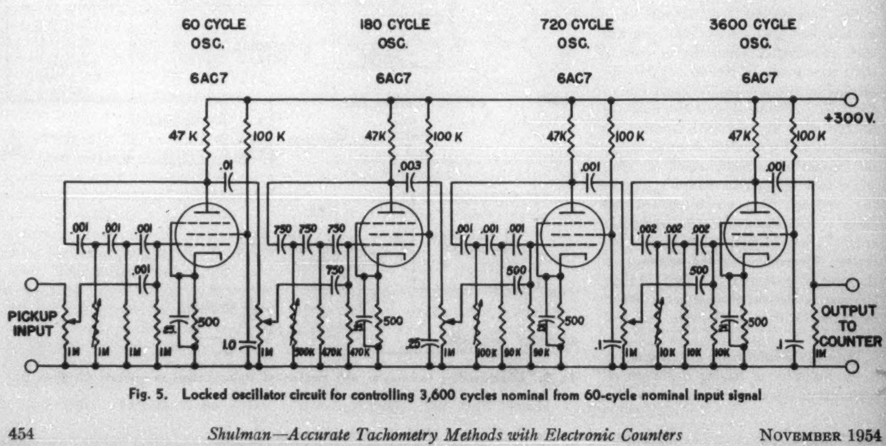

60 CYCLE OSC.
180 CYCLE OSC.
720 CYCLE OSC.
3600 CYCLE OSC.
6AC7
6AC7
6AC7
6AC7
+300V.
47 K
100 K
47K
100 K
47K
100 K
47K
100 K
.01
.003
.001
.001
.001 .001 .001
750 750 750
.001 .001 .001
.002 .002 .002
.001
750
500
500
PICKUP INPUT
OUTPUT TO COUNTER
500
500
500
500
25
1.0
.25
.1
.1
1M
500K
470K
470K
100K
90K
90K
10K
10K
10K

Fig. 5. Locked oscillator circuit for controlling 3,600 cycles nominal from 60-cycle nominal input signal

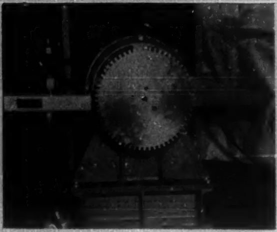


Fig. 6 (left). Magnetic pickup and 60-tooth disk signal generator combination

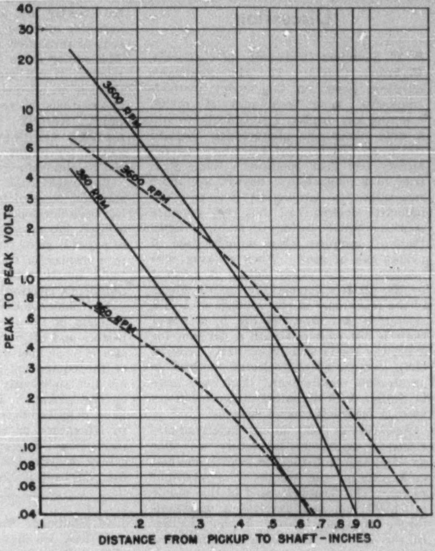


Fig. 7 (right). Pickup response characteristics: Solid line, with disk; dash line, with keyway in shaft

three tachometry methods described. Photoelectric pickups of several types were tried on the test floor. Because they could not stand abuse and were difficult to adjust and susceptible to stray pickup, none of these were considered entirely satisfactory. A rugged magnetic pickup was designed for use with a 60-tooth disk, as shown in Fig. 6. This pickup has enough sensitivity to be used 1/2 inch or more from either the disk or a shaft, as shown in Fig. 7. The shape of wave forms obtained in each case is shown in Fig. 8.

Summary

From the standpoint of operating convenience, the electronic tachometer shows real worth on a production test floor, where it is usually impractical to make and use a laboratory-type setup for getting accurate speed measurements. The counter portion of the electronic tachometer may be mounted on or near the meter board in such a way that the tester can take speed readings along with other meter readings. A second operator is not required to hold or watch the pickup, which can be mounted on a tripod adjacent to the shaft and left alone. A coaxial cable suffices to connect the pickup with the counter system. In the first and third methods described, the pickup can be moved quickly from one shaft to another. Except for phase reversal the signal generated by the pickup is the same whether the keyway is open or filled with a steel key; therefore, speed can be measured whether or not the motor is coupled to a load.

References

1. Electronic Tachometer, W. Richter. *Machine Design*, Cleveland, Ohio, vol. 21, no. 11, Nov. 1949, pp. 133–36.

2. Precision Speed Measurement of Rotating Equipment, M. W. Hellar. *General Electric Review*, Schenectady, N. Y., vol. 52, no. 10, Oct. 1949, pp. 22–26.

3. Precision Measurement of Rotary Motion, H. J. Finden. *Electronic Engineering*, London, England, vol. 22, no. 263, Jan. 1950, pp. 2–8.

4. A Recording Tachometer for Measuring Instantaneous Angular Speed Variations, S. P. Bartles. *Electrical Engineering*, vol. 70, Sept. 1951, pp. 816–19.

5. A New High-Accuracy Counter-Type Tachometer, T. M. Berry, C. L. Beattie. *Electrical Engineering*, vol. 69, July 1950, p. 605.

6. An Electronic Tachometer, H. G. Jerrard, S. W. Punnett. *Journal of Scientific Instruments*, London, England, vol. 27, Sept. 1950, pp. 244–45.

7. A New 100 KC Counter for Use in Electronics and Industry, E. A. Hilton. *Hewlett Packard Journal*, Palo Alto, Calif., vol. 4, no. 3, Nov. 1952.

8. Vacuum Tube Oscillators (book), W. A. Edson. John Wiley & Sons, Inc., New York, N. Y., 1953, chapters 13 and 14.

9. Phase Shift Oscillators, E. L. Ginzton, L. M. Hollingsworth. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 29, Feb. 1941, pp. 43–40.

10. Phase Shift Oscillator Design Charts, W. W. Kunde. *Electronics*, New York, N. Y., Nov. 1943, pp. 132–33.

11. Extending the Frequency Range of the Phase Shift Oscillator, R. W. Johnson. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 33, Sept. 1945, pp. 597–603.

12. The Tapered Phase Shift Oscillator, P. G. Sulzer. *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 36, Oct. 1948, pp. 1302–05.

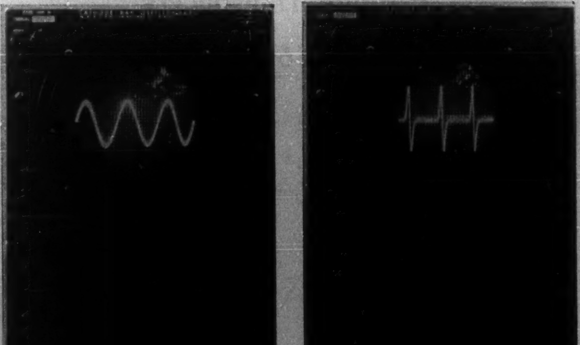


Fig. 8. Signal wave forms from pickup: (left) with disk; (right) with shaft keyway

Discussion

R. M. Saunders (University of California, Berkeley, Calif.): To Mr. Shulman's excellent paper on tachometry methods employing some techniques of digital counting and display, I should like to add a note about the photoelectric method which he dismisses as being an unsatisfactory test floor operation. While it is true that photoelectric devices are more susceptible to mechanical damage than inductive devices, and that they generate spurious signals owing to ambient light, there are instances where no other type of pickup can be used. When working with machines whose power output is affected by the variable magnetic reluctance effects of the transducer or the windage of a disk similar to that shown in Fig. 6, or whose shaft is not equipped with a flat-spot for a keyway (or in some cases never brought out), recourse to photoelectric methods is an absolute requirement. Such cases arise in dealing with actuators and transducers used in feedback control systems.

As a case in point, in our undergraduate electrical machinery laboratory we offer an experiment on the acceleration method of measuring the torque of a gyrospin motor. The terminal speed of this motor is 24,000 rpm and develops a maximum torque of approximately 5 ounce-inches. In this case the tachometry is performed by cutting a small window in the housing of the gyro and inserting a clear plastic window. The rotor is then painted with six alternate black and white segments which are looked at by a photoelectric light source and cell combination. The output from the photoelectric cell is then fed directly into the events-per-unit-time indicator and the speed is then given in tens of revolution on a 4-register counter. This motor takes 5 to 10 minutes to accelerate to full speed. During that time there is an adequate period in which 100 observations of the speed of the motor can be recorded. By subtracting adjacent readings, and knowing the moment of inertia, the torque developed by the motor can be obtained.

For application to larger motors, a standard strip of paper with 60 segments has been developed for use with our standardized laboratory couplings. This strip of paper is glued on the couplings resulting in a display in rpm, thus performing the same function of multiplication as Mr. Shulman's electronic multiplier.

Mr. Shulman has very nicely solved the problem of the testing of large induction motors and other machines of fairly large size on a test floor where, by the very nature of the equipment involved, the workmen are not sufficiently careful to use photoelectric pickups. The situations which I have described have been applied largely to laboratory or testing precision equipment, where the care exercised is more the caliber of instrument work, and the danger from mechanical shock is greatly minimized.

J. M. Shulman: As Mr. Saunders points out, there are many applications where a photoelectric pickup system might be better than a magnetic pickup, or where a magnetic pickup might be ruled out entirely. The remarks about pickups were not intended to discourage the use of photoelectric pickups where they are applicable.

Mr. Saunders' comments on how he is using the counter to obtain speed-torque characteristics on a motor having long starting time are of interest. Accurate speed-torque curves on large motors are difficult to obtain by load testing, and the so-called inertia method in which angular acceleration is measured as a function of time and plotted as a function of speed during the starting period is potentially a solution to this problem. The method used by Mr. Saunders is of course much too slow for most motors, which start in a few seconds or a fraction of a second. Considerable work, notably that of S. S. L. Chang of New York University, has been done in making speed-torque curves by photographing an oscilloscope trace, with the use of electronic means to obtain speed as a function of time and to differentiate the speed-time characteristic. H. W. Hansen of Westinghouse Electric Corp., Sunnyvale, Calif. and the author have developed a number of improvements in this method and hope to present the results in the near future.

John Corl of Berkeley Scientific Division, Beckman Instruments Inc., Richmond, Calif., has pointed out that any system using free-running oscillators is not fail-safe. This is true and must be taken into account when using the third method. Mr. Corl proposes that a fail-safe direct-reading tachometer could be formed by a digital converter which would measure the time interval, as described in the first method, and operate on this number by taking the reciprocal and multiplying by 6×10^n, where n is determined by the unit of time used in the measurement. If this could be done at reasonable cost, it would form an ideal tachometer from the standpoint of the requirements listed in the paper.

Networks for Digital-to-Analogue Shaft-Position Transducers

S. J. O'NEIL
ASSOCIATE MEMBER AIEE

DIGITAL computers have been used extensively in the solution of mathematical problems requiring both speed and accuracy in their computation. More recently, digital computers are being employed to control analogue equipment. The control of machine tools is one example of this application.[1]

In the solution of mathematical problems it is often sufficient to present the output in the form of a table of numbers. In the control of analogue equipment, however, the digital computer output must first be converted to analogue form in order to vary some analogue input to the controlled equipment.

The analogue-to-digital conversion problem is as old as the digital computer. It is only recently that the inverse problem (the digital-to-analogue conversion) has become important. This paper deals with the networks required in a particular type of digital-to-analogue shaft-position converter.

The converters (or transducers) for which these networks are designed have been described previously.[2] Alternating or direct voltages which are approximate sine functions of the digital input are applied to the stator windings of 2-phase or 3-phase electromechanical machines, such as resolvers or synchros. Approximate sine waves composed of straight-line segments are used to reduce the number of switching components necessary. The direction of the magnetic field induced by the stator voltages is compared with the direction of the rotor coil axis. The rotor shaft turns through an angle proportional to the digital input. Magnetic converters of this type are inherently minor-arc sensing devices.

The Problem

The problem is to design networks which produce voltages which are approximate 2-phase or 3-phase sine functions of a binary digital number. The

---

Paper 54-272, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 11, 1954; made available for printing April 27, 1954.

S. J. O'Neil is with the Parke Mathematical Laboratories, Inc., Concord, Mass.

This work was accomplished at the Air Force Cambridge Research Center, with which the author was formerly associated. The author is indebted to R. P. Bigliano for the design of the basic network for obtaining linear voltage wave forms.

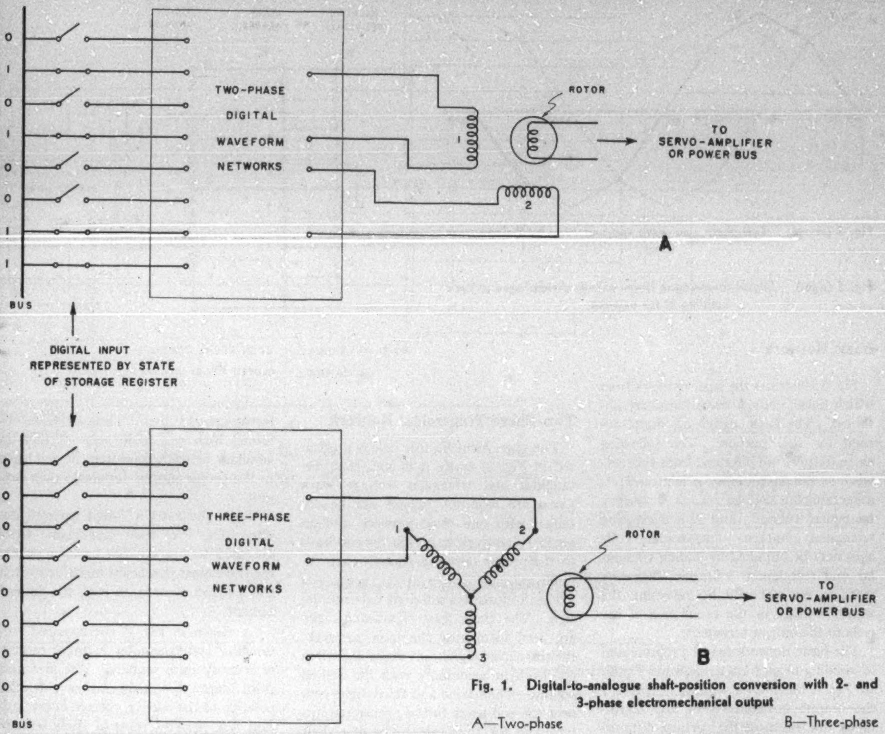

0
1
0
1
0
0
1
1
BUS
TWO-PHASE
DIGITAL
WAVEFORM
NETWORKS
1
2
ROTOR
TO
SERVO-AMPLIFIER
OR POWER BUS
A
DIGITAL INPUT
REPRESENTED BY STATE
OF STORAGE REGISTER
0
0
0
0
0
0
0
0
BUS
THREE-PHASE
DIGITAL
WAVEFORM
NETWORKS
1
2
3
ROTOR
TO
SERVO-AMPLIFIER
OR POWER BUS
B

Fig. 1. Digital-to-analogue shaft-position conversion with 2- and 3-phase electromechanical output

A—Two-phase B—Three-phase

digital number represents the state of a parallel bank of relays. The output voltages are applied to the stators of resolver or synchro machines which are used with or without servo-amplification to establish a shaft position proportional to the binary digital number. The position of the desired network between the input storage register and the 2-phase or 3-phase stator output is illustrated in Fig. 1.

Desired Wave Forms

The desired wave forms are illustrated in Figs. 2, 3, and 4. In these figures the sine waves are approximated by sections of straight lines. The approximate wave forms are continuous with discontinuities in slope or corner points occuring periodically every octant (for the 2-phase wave forms) or sextant (for the 3-phase wave forms). In the intervals between successive corner points, the voltages are either constant or they vary linearly with an increase in the digital input. During these intervals rectangular, triangular, and trapezoidal voltage wave forms are needed. As in the case of actual sine or cosine waves the approximate wave forms are odd or even functions depending upon the position of the origin.

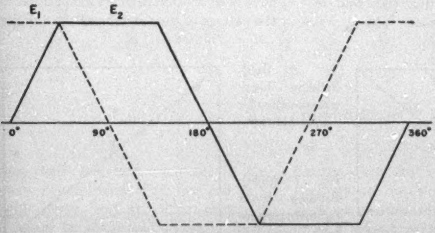


Fig. 2. Two-phase sine wave approximation by trapezoidal wave forms

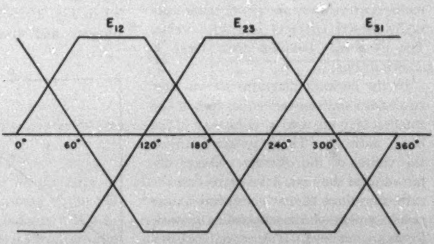


Fig. 3. Three-phase sine wave approximation by trapezoidal wave forms

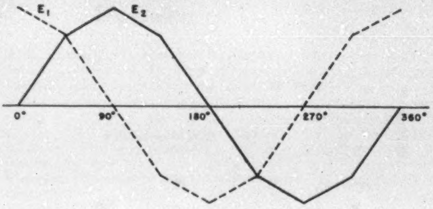


Fig. 4 (above). Two-phase sine wave approximation by 2-slope wave forms

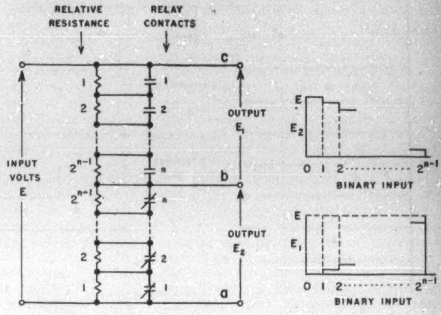


Fig. 5 (right). Digital-to-analogue linear voltage divider used as basic building block network

Basic Network

Fig. 5 illustrates the basic network from which linear voltage wave forms are obtained. The least significant digits are used for this purpose. The following three outputs are obtained from this network as the digital input is increased: 1. a rectangular voltage; 2. an increasing triangular voltage; and 3. a decreasing triangular voltage. Trapezoidal voltages may be obtained by adding triangular and rectangular voltages. Negative voltages are obtained by reversing the supply voltage or the connection of the coils to the output terminals.

The basic network may be represented in building-block form as shown in Fig. 6. In Fig. 6 the lines *a*, *b*, and *c* represent the output terminals. For any digital input the length of the vertical distance between the lines *a*, *b*, and *c* gives the relative magnitude of the voltage output between these terminals.

In some situations it is necessary to reverse the slope of the triangular voltage applied to a coil. This can be accomplished either by switching the coil to a different set of output terminals or by reversing the slope of the voltage output at the original set of terminals. The slopes of the voltage wave forms from the basic network may be reversed if all the normally open relay contacts are made normally closed contacts and vice versa. The resulting building-block form is shown in Fig. 7.

In the network diagrams shown later two additional resistors equal to half the smallest step are used at either end of the basic network. These resistors adjust the values of the discrete voltages obtained from the basic network so that the extreme values which correspond to the corner points of the approximate wave forms are not used. This procedure helps to ensure the uniqueness of the resulting shaft-position angles.

Two-Phase Trapezoidal Network

The wave form for this case is illustrated in Fig. 2, where it is seen that rectangular and triangular voltage wave forms are required. These can be obtained with one basic network and an auxiliary network to switch the coil leads to the proper output terminals each octant.

Fig. 8 illustrates a 6-digit network design. The three least significant digits are used to operate the basic network; the three most significant digits, to switch coil leads in accordance with the desired octant. The second and third digits connect the coil leads to the proper output-terminals for 1/2 cycle. The first digit reverses the polarity of the power supply voltage every half-cycle to give a mirror image of the wave form for the first half-cycle reflected about the zero axis. Alternatively, instead of reversing the polarity of the power supply, the first digit could be used to reverse the polarity of the rotor coil.

Three-Phase Trapezoidal Network

The wave form for this case is illustrated in Fig. 3 where it is seen that rectangular and triangular voltage wave forms are required. These can be obtained with one basic network and an auxiliary network to switch the coil leads to the proper output terminals each sextant.

Fig. 9 illustrates a 7-digit network design. The four least significant digits are used to operate the basic network; the three most significant digits, to switch coil leads in accordance with the desired sextant.

As shown in Fig. 3 the required rectangular and triangular voltages reverse in polarity each sextant. The first and third digits are used to reverse the polarity of the supply voltage every sextant. A polarity reversal each sextant causes the supply voltage during 1/2 cycle to be opposite to that during the next half-cycle. Therefore, the switching circuit for connecting the coil leads to the proper output terminals needs to be designed for 1/2 cycle only. The second and third digits are used for this purpose.

Since only six of the eight possible combinations of the first three digits are necessary to select the proper sextant, two have been omittted. In the network of Fig. 9 the two combinations which have been omitted result in a repetition of the pattern of the second and third digits

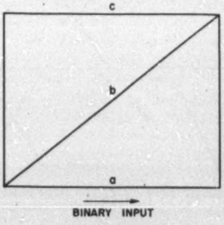


Fig. 6 (left). Building-block representation for basic network

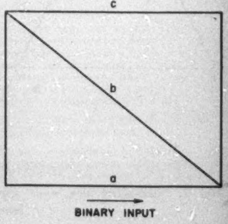
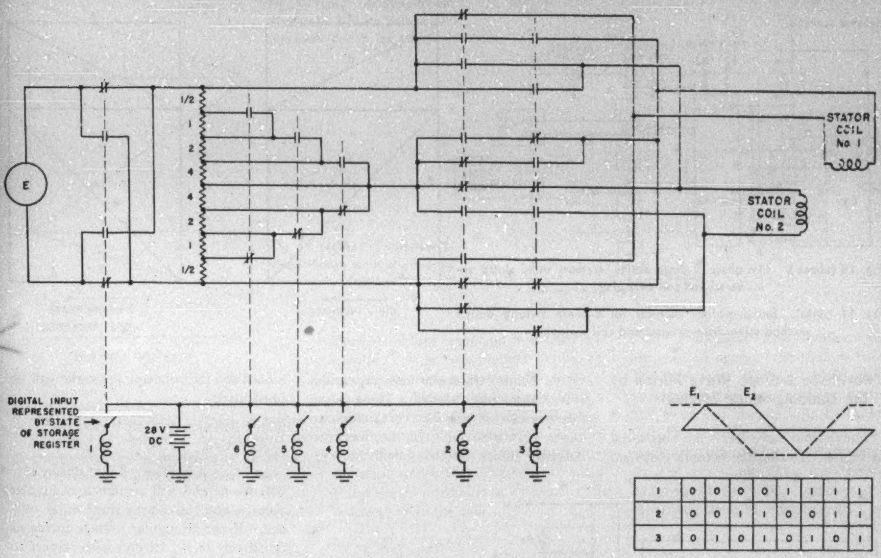
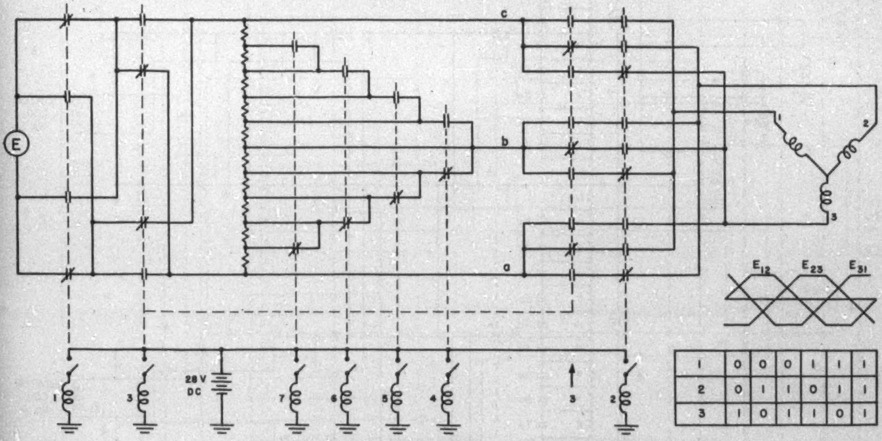


Fig. 7 (right). Building-block representation for basic network with slope reversed


1/2
1
2
4
4
2
1
1/2
E
STATOR COIL No. 1
STATOR COIL No. 2
E₁
E₂
DIGITAL INPUT REPRESENTED BY STATE OF STORAGE REGISTER
28 V DC
1
6
5
4
2
3

1	0	0	0	0	1	1	1	1
2	0	0	1	1	0	0	1	1
3	0	1	0	1	0	1	0	1

Fig. 8. Network diagram for 2-phase trapezoidal wave forms


c
b
a
E
1
2
3
E₁₂
E₂₃
E₃₁
28 V DC
1
3
7
6
5
4
3
2

1	0	0	0	1	1	1
2	0	1	1	0	1	1
3	1	0	1	1	0	1

Fig. 9. Network diagram for 3-phase trapezoidal wave forms

every half-cycle and a change in the binary noncarry sum of the first and third digits every sextant. This choice is necessary for the proper operation of the network of Fig. 9.

Since not all of the possible digit combinations are used, only 96 divisions of the circle are obtained instead of the 128 divisions possible with 7 digits. The 3-phase synchro introduces a three into the binary problem so that either an inefficient binary coding method and a simple network design, as shown here, or a much more complicated switching network is necessary. Two-phase machines lead to a natural division of the circle into binary parts and can result in an efficient coding method and a simple network design.

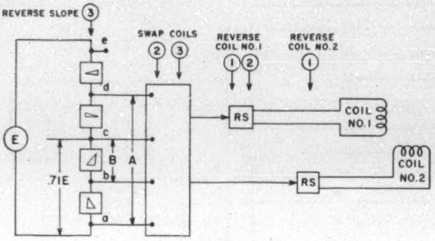

REVERSE SLOPE ③
SWAP COILS ② ③
REVERSE COIL NO. 1 ① ②
REVERSE COIL NO. 2 ①
e
d
c
b
a
E
B
A
.71E
RS
COIL NO. 1
RS
COIL NO. 2

Fig. 10 (above). Two-phase 2-slope design method using slope reversal and coil swapping

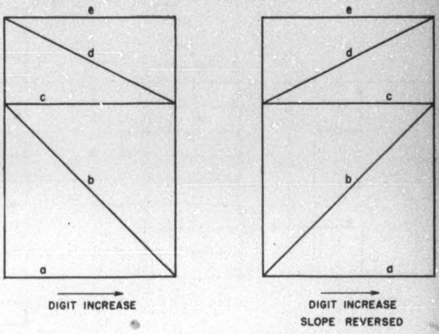

e
d
c
b
a
DIGIT INCREASE
e
d
c
b
a
DIGIT INCREASE
SLOPE REVERSED

Fig. 11 (right). Building-block diagram for 2-phase 2-slope design method using slope reversal and coil swapping

Two-Phase 2-Slope Wave Forms by the Building-Block Method

The desired wave forms are illustrated in Fig. 4. During the octants (between corner points) triangular and trapezoidal wave forms are needed. These wave forms can be obtained by combining two basic networks. In this section two different design methods will be presented and the resulting networks will be illustrated.

Slope-Reversing Coil-Swapping Network

In Fig. 4 it is observed that during any octant one coil will require a triangular voltage, and the other a trapezoidal voltage. If the triangular voltage increases (positively or negatively) with respect to

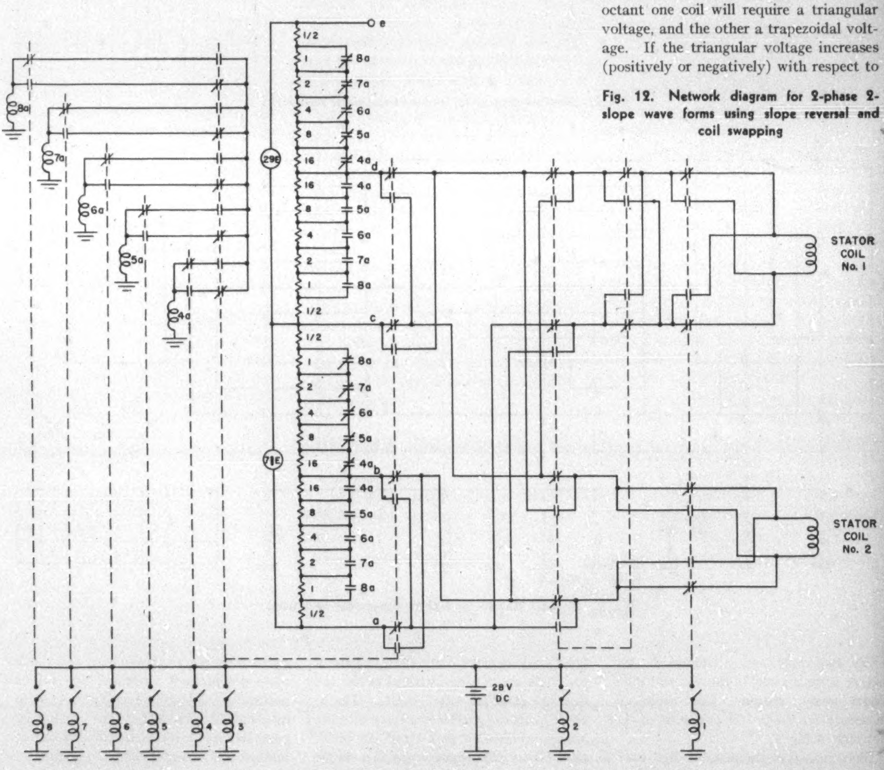

e
d
c
b
a
1/2
1
2
4
8
16
8a
7a
6a
5a
4a
29E
71E
STATOR COIL No. 1
STATOR COIL No. 2
28 V DC
8
7
6
5
4
3
2
1

Fig. 12. Network diagram for 2-phase 2-slope wave forms using slope reversal and coil swapping

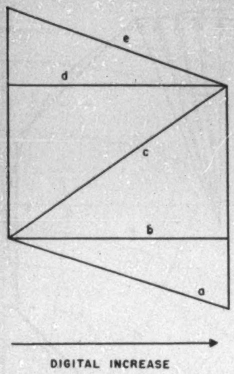

e
d
c
b
a
DIGITAL INCREASE

Fig. 13 (left). Building-block diagram for 2-phase 2-slope design method using tap selection

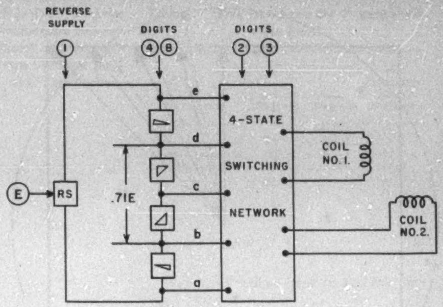

REVERSE SUPPLY
DIGITS
DIGITS
①
④⑧
②③
e
d
c
b
a
4-STATE
SWITCHING
NETWORK
COIL NO.1.
COIL NO.2.
E
RS
.71E

Fig. 14 (right). Two-phase 2-slope design method using tap selection

zero, the trapezoidal voltage decreases. The type of change in the triangular or trapezoidal wave (increase or decrease) with respect to zero alternates each octant.

Fig. 10 illustrates a design method. During the first octant an increasing triangular voltage is obtained across *bc*, and a decreasing trapezoidal voltage, across *ad*. During subsequent octants a decreasing triangular voltage and an increasing trapezoidal voltage are obtained by reversing the slope of the output from the basic networks. The effect of negative voltages is obtained by reversing the connections to the coils.

In Fig. 10 the five least significant digits are used to operate the basic networks. The three most significant digits are used to reverse slopes, reverse coil connections, and swap coil circuits. As indicated in Fig. 10 the voltages required at the coil leads are obtained by: 1. Reversing the slope of the outputs of the basic networks each octant (with digit number 3); 2. Reversing the connections to coil 2 every

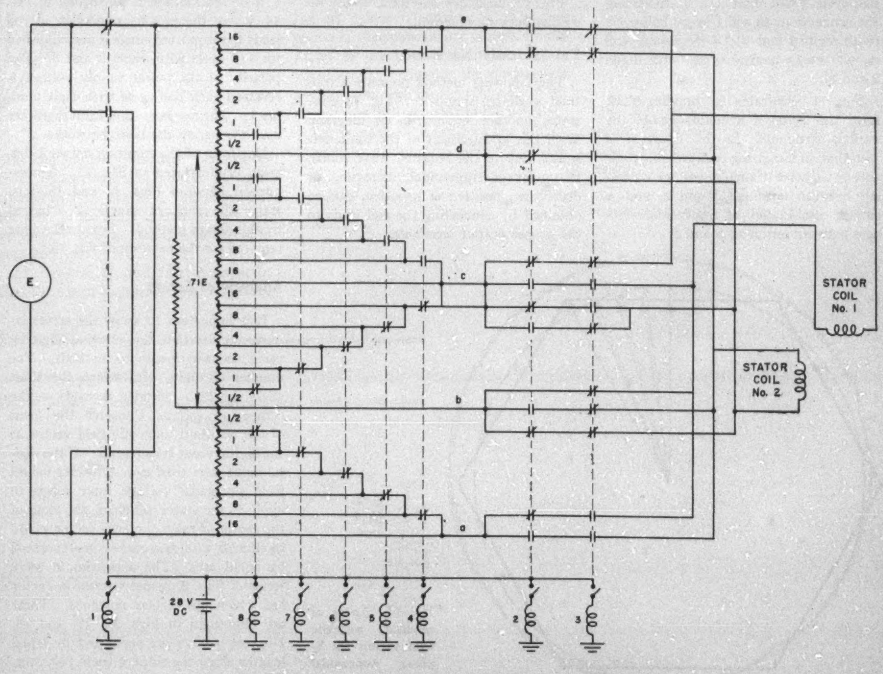

16
8
4
2
1
1/2
1/2
1
2
4
8
16
16
8
4
2
1
1/2
1/2
1
2
4
8
16
e
d
c
b
a
.71E
E
STATOR COIL No. 1
STATOR COIL No. 2
28 V DC
1
8
7
6
5
4
2
3

Fig. 15. Network diagram for 2-phase 2-slope wave forms using tap selection

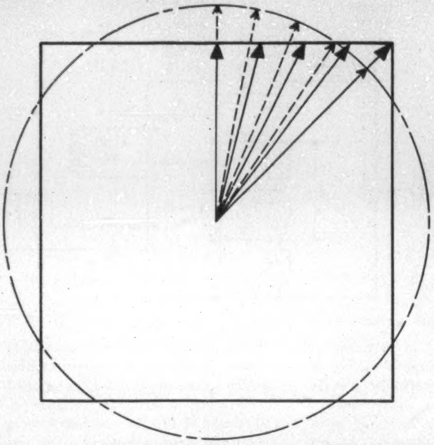


Fig. 16. Locus of resultant magnetic field vectors for 2-phase trapezoidal wave forms

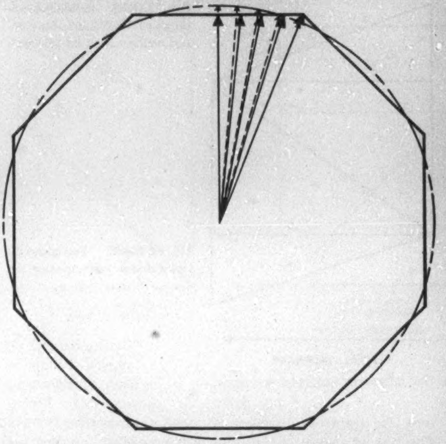


Fig. 18. Locus of resultant magnetic field vectors for 2-phase 2-slope wave forms

half-cycle (with digit 1); 3. Reversing the connections to coil 1 every half-cycle (with digits 1 and 2); 4. Swapping coil circuits every quarter-cycle (with digits 2 and 3).

Fig. 11 illustrates in building-block form the voltages available across the various terminals. In this figure it is seen that an increasing or decreasing positive or negative triangular voltage is present between terminals *b* and *c*, and a similar combination of trapezoidal voltages between terminals *a* and *d*.

Fig. 12 illustrates one form of the resulting network diagram.

Tap Selection Network

The following method presents a contrast in design simplicity: Fig. 13 illustrates another building-block diagram. With this arrangement of the basic networks, any of the required wave forms, triangular or trapezoidal, increasing or decreasing, positive or negative, can be obtained by connecting the coil leads to the proper output terminals.

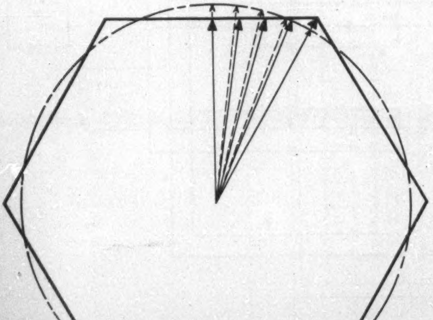


Fig. 17. Locus of resultant magnetic field vectors for 3-phase trapezoidal wave forms

A design method is illustrated in Fig. 14, where the proper connection of coil leads to output terminals is accomplished for 1/2 cycle with digits 2 and 3. The polarity of the power supply voltage is reversed each half-cycle with digit number 1. The five least significant digits are used to operate the basic networks.

One form of the resulting network diagram is illustrated in Fig. 15. A comparison between Figs. 12 and 15 illustrates the network simplicity achieved by this design method. Fewer relays are required for the network of Fig. 15.

Analysis of Errors

It is important to know the errors introduced into the shaft-position angle by using these approximate methods. The number of digits with which the shaft angle should be encoded depends on the conversion errors. Consider the locus of the resultant magnetic field vector as the digital input is increased. If the digital input were used to select exact values from sinusoidal voltage wave forms to apply to the stator windings, the locus of the resultant magnetic field vector would be circular with discrete vectors separated by equal arcs. The approximate wave forms of Figs. 2, 3, and 4 result in vector loci which are regular polygons. These are illustrated in Figs. 16, 17, and 18. Discrete vectors are separated by equal lengths along the sides of these polygons.

It is obvious from these illustrations

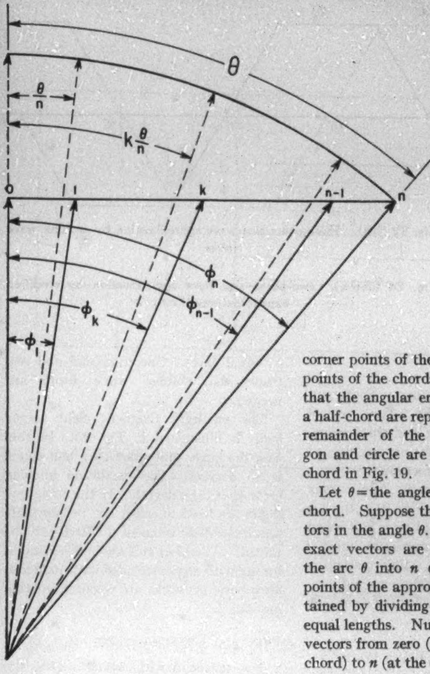


Fig. 19. Angular relationship between exact and approximate resultant magnetic field vectors

that the method presented here essentially approximates a number of equal circular arcs by chords. The analysis of the error in the angle associated with any vector now becomes a simple problem in geometry.

In Figs. 16, 17, and 18 it is apparent that there is no angular error at either the corner points of the polygons or the midpoints of the chords. It is also apparent that the angular errors which occur over a half-chord are repetitive throughout the remainder of the polygon. The polygon and circle are illustrated for a half-chord in Fig. 19.

Let θ = the angle subtended by a half chord. Suppose there are n discrete vectors in the angle θ. The positions of the exact vectors are obtained by dividing the arc θ into n equal arcs. The end points of the approximate vectors are obtained by dividing the half chord into n equal lengths. Number the approximate vectors from zero (at the midpoint of the chord) to n (at the corner of the polygon).

Let ϕ_k = the angle between the kth approximate vector and the zero vector. Then

$$\phi_n = \theta \tag{1}$$

Also

Table I. Errors Introduced by Approximate Wave Forms

Wave Form	θ, Degrees	$\epsilon_{\max}$, Degrees	$\phi\|$ Maximum Error, Degrees
2-phase trapezoid	45	±4.1	27.6
3-phase trapezoid	30	±1.1	17.8
2-phase 2-slope	22.5	±0.4	13.2

$$\tan \phi_k = \frac{k}{n} \tan \theta \tag{2}$$

Let ϵ_k = angular error in the kth vector. Then

$$\epsilon_k = \frac{k}{n}\theta - \tan^{-1}\left(\frac{k}{n}\tan\theta\right) \tag{3}$$

If equation 3 is differentiated with respect to k and then equated to zero, the following result is obtained

$$\frac{k}{n} = \sqrt{\frac{1}{\theta \tan \theta} - \frac{1}{\tan^2 \theta}} \tag{4}$$

The maximum angular error is obtained by substituting equation 4 into equation 3 with the following result

$$\epsilon_{\max} = \sqrt{\frac{\theta}{\tan\theta} - \frac{\theta^2}{\tan^2\theta}} - \tan^{-1} \sqrt{\frac{\tan\theta}{\theta} - 1} \tag{5}$$

The angle at which the maximum error occurs is obtained by substituting equation 4 into equation 2 with the following result

$$\tan \phi\big|_{\text{max error}} = \sqrt{\frac{\tan\theta}{\theta} - 1} \tag{6}$$

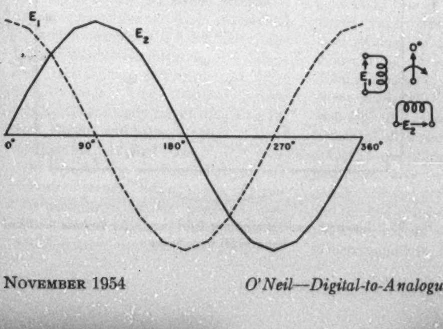


Fig. 20 (below). Two-phase sine wave approximation by 4-slope wave forms

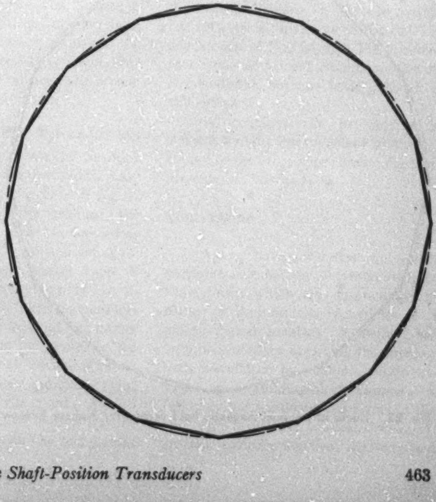
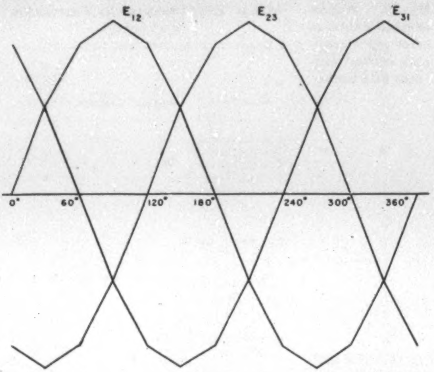


Fig. 21 (right). Locus of resultant magnetic field vectors for 2-phase 4-slope wave forms



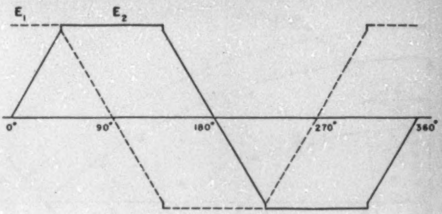


Fig. 22 (left). Three-phase sine wave approximation by 3-slope wave forms

Fig. 24 (above). Two-phase sine wave approximation by modified trapezoidal wave forms

The results for the three wave forms considered in this paper are summarized in Table I.

Possible Extension of Method

TWO-PHASE 4-SLOPE WAVE FORM

Since a circle is approximated more closely by increasing the number of chords, it is reasonable to expect that the angular error in any approximate vector will decrease as the number of chords is increased. The results of the previous section show that this is the case.

A logical extension of the methods presented in this paper in the 2-phase case is illustrated by the wave forms of Fig. 20, where the period is divided into 16 equal parts. Twice as many linear sections are used as in the 2-slope case of Fig. 4. One triangular and three trapezoidal voltage wave forms are needed during the various semioctants.

The resultant magnetic field vector locus is illustrated in Fig. 21. In this case the angle subtended by a half-chord is 11.25 degrees. The maximum angular error is ±0.04 degrees. This method is suitable for use with angles encoded with 12 digits. A more complicated switching network is necessary.

THREE-PHASE 3-SLOPE WAVE FORM

A more accurate method in the 3-phase case is illustrated by the wave forms of Fig. 22, where the period is divided into 12 equal parts. One triangular and two trapezoidal voltage wave forms are needed.

The resultant magnetic field vector locus is illustrated in Fig. 23. In this case the angle subtended by a half-chord is 15 degrees. The maximum angular error is ±0.13 degree. If the first four digits are used to select the proper half-sextant, 768 divisions of the circle are obtained. Therefore this method is suitable for use with angles encoded with 10 digits. Two basic networks are required for this method.

PRACTICAL CONSIDERATIONS

The precision with which a quantity can be measured or reproduced is given by the square root of the sum of the squares of all the independent errors present.

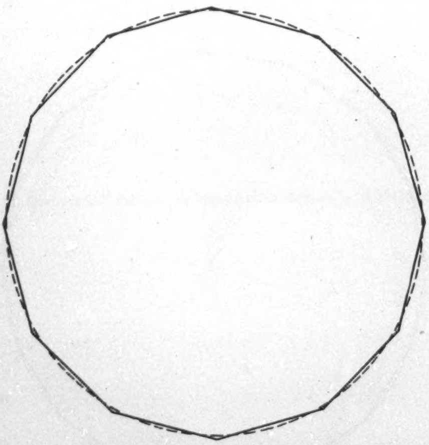


Fig. 23. Locus of resultant magnetic field vectors for 3-phase 3-slope wave forms

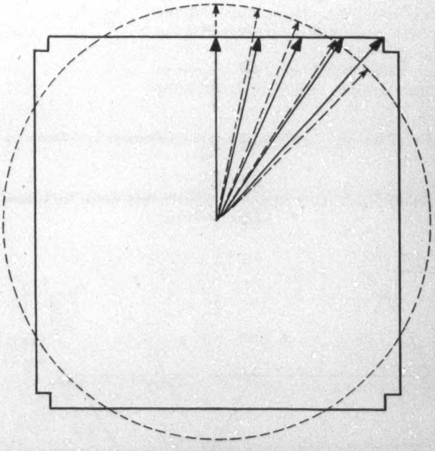
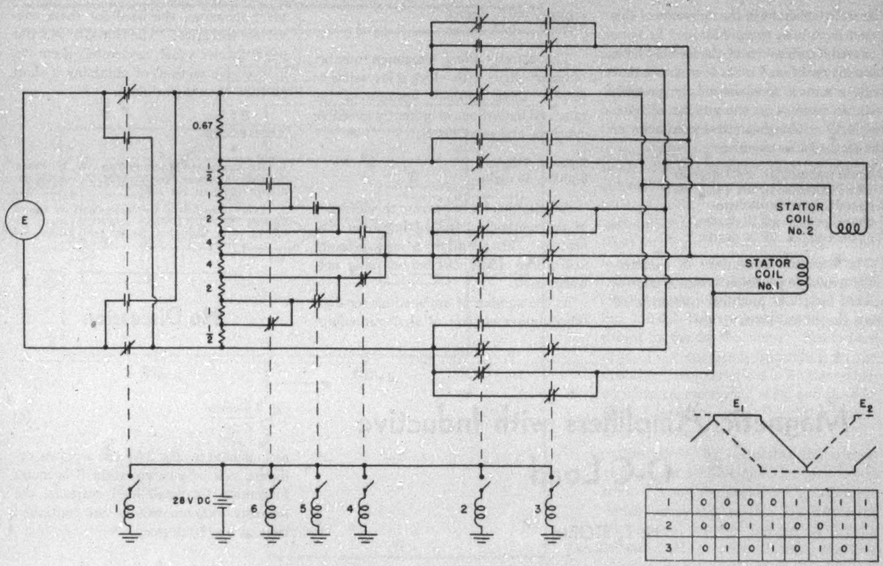


Fig. 25. Locus of resultant magnetic field vectors for 2-phase modified trapezoidal wave forms



1	0	0	0	0	1	1	1	1
2	0	0	1	1	0	0	1	1
3	0	1	0	1	0	1	0	1

Fig. 26. Network diagram for 2-phase modified trapezoidal wave forms

Some of the sources of error present are: 1. the electromechanical machine; 2. the approximate wave forms; 3. the components in the switching circuitry; 4. the angular quantizing interval. The angular quantizing interval is a measure of the accuracy desired in the output shaft position. So that all the shaft-position digits be significant, the precision of the converter in reproducing any of the discrete angular input should be of the order of half the quantizing interval. Greater precision, if attainable without additional complexity and cost, is desirable but not necessary.

The precision inherent in the wave forms for a method determines the theoretical usefulness of the method. From the results shown in previous sections it appears that the 2-phase trapezoidal method is suitable for a 6-digit code; the 3-phase trapezoidal method, for a 7-digit code; the 2-phase 2-slope method, for a 9-digit code. The 3-phase 3-slope method theoretically can be used for a 10-digit code, and the 2-phase 4-slope method, for a 12-digit code.

At present the electromechanical machines represent a limiting source of error. Commercially available precision synchros are accurate to approximately ±0.25 degree; resolvers, to ±0.67 degree. If synchros or resolvers are used as the electromechanical members in these converters, the practical limitations on the digital accuracy attainable are ten digits for the 3-phase methods, and eight digits for the 2-phase methods.

These considerations rule out the practical use of the 2-phase 4-slope method with present resolvers. The 2-phase 2-slope method used with eight binary digits utilizes the full accuracy of present resolvers. The 3-phase 3-slope method can be used to obtain the full accuracy of commercially available synchros.

Reduction of Errors

It is possible to increase the accuracy of the various methods presented by making a slight change in the shape of the assumed wave forms. For example, the wave form of Fig. 2 might be modified as shown in Fig. 24. The corresponding resultant magnetic field vector locus is shown in Fig. 25. By comparing Fig. 25 with Fig. 16 it is seen that this modification introduces some error at the corner points of the polygon but decreases the maximum error, which occurs approximately halfway between the corner point and the mid-point of the chord. A trial-and-error method can be quickly applied to determine how much the rectangular voltage should be increased to make the error introduced at the corner point equal to the resulting maximum error.

With this technique the maximum error for the 2-phase trapezoidal method can be reduced from ±4.1 to ±2.3 degrees. This will make the 2-phase trapezoidal method more suitable for use with 6-digit angular codes. A suitable switching network is illustrated in Fig. 26. A comparison of Fig. 26 with Fig. 8 shows that some additional circuit complexity is introduced, but one basic network is still sufficient.

This technique can be applied to achieve a slight improvement in the accuracy of the other approximate methods presented in this paper.

Conclusions

This paper has presented several simple networks suitable for obtaining voltage wave forms which are approximate 2-phase or 3-phase sine functions of a binary digital number. A method has been presented for obtaining the approximate wave forms by combining basic networks in a building-block manner. A technique for analyzing the errors in the approximate methods has been developed; an error equation has been derived; and the errors inherent in the networks of this paper have been computed.

Possible extensions of the method have been discussed, and it has been shown that high accuracy is theoretically possible with an increase in the number of basic networks. The theoretical accuracies of the methods discussed are:

2-phase trapezoidal, ±4.1 degrees
3-phase trapezoidal, ±1.1 degrees
2-phase 2-slope, ±0.4 degree
3-phase 3-slope, ±0.13 degree
2-phase 4-slope, ±0.04 degree

The accuracy limitations of commercially available precision synchros and resolvers limit the practical usefulness of these theoretical methods to:

2-phase, ±0.67 degree
3-phase, ±0.25 degree

The corresponding maximum number of binary digits with which it is feasible to employ these methods because of the practical limitations in presently available synchros and resolvers is:

2-phase, 8 digits
3-phase, 10 digits

A technique for improving the accuracy of the methods discussed has been presented. This results in a more efficient utilization of the number of basic networks used.

As the number of applications for digital-computer control of analogue equipment increases, the need for these converters will grow. The converters of this paper furnish a fast, moderately accurate, inexpensive method of obtaining a shaft position from digital data.

References

1. AN AUTOMATIC MACHINE TOOL, W. M. Pease. *Scientific American*, New York, N. Y., vol. 187, no. 3, Sept. 1952, pp. 101–15.

2. DIGITAL-TO-ANALOGUE SHAFT-POSITION TRANSDUCERS, S. J. O'Neil. *AIEE Transactions* vol. 72, pt. I, Mar. 1953, pp. 37–41.

◆

No Discussion

Magnetic Amplifiers with Inductive D-C Load

H. F. STORM
MEMBER AIEE

Synopsis: The occurrence of instabilities in magnetic amplifiers with inductive d-c load is well known. Less known, however, is the mechanism causing the instability, of which a published analysis is believed nonexistent. The purpose of this paper is to discuss the mechanics of instability qualitatively and quantitatively and to explain some of the means by which this instability can be averted.

THE circuit of the magnetic amplifier is shown in Fig. 1. The two saturable reactors A and B are energized from the secondary TS of a supply transformer; terminals $X2$ of the gate windings lead to the rectifiers REC_1, REC_2, which, in turn connect to the load. The load consists of resistance R_L and a linear inductance L_L. The current in the load is determined by the control voltage E_C, which is the independent variable. If $L_L=0$, the load becomes purely resistive, and the control characteristic assumes the typical cosine-like shape as shown in Fig. 2 ($\omega L_L/R_L=0$) and discussed in reference 1.

The per-unit load current $\mathbf{i}_L$ in Fig. 2 is

$$\mathbf{i}_L=\frac{I_L}{I_{L,m}} \tag{1}$$

$$I_{L,m}=\frac{E}{R_L} \tag{2}$$

where

E = average value of the rectified supply voltage e_{1-N} or e_{2-N}
R_L = load resistance into which gate and rectifier forward resistances have been lumped

On the abscissa axis are plotted the per-unit control ampere-turns

$$\mathbf{a}=\frac{I_C N_C}{I_{G,m} N_G}=\frac{I_C N_C}{I_{L,m}(N_G/2)} \tag{3}$$

$I_{L,m}(N_G/2)$ being the ampere-turns of one gate averaged over an entire cycle. The abscissa for the minimum of $\mathbf{i}_L$ (point M) is $-I_X/I_{L,m}$, where I_X is the average value of the exciting current of the saturable reactor with both gate windings connected in parallel (reference 1, I_E corresponds to I_X) and with control windings open-circuited. The supply voltages of e_{1-N}, e_{2-N} are sinusoidal and of such magnitude and frequency as to produce a change of flux density from $-B_s$ to B_s, when applied to the gates of the saturable reactors A and B.

With saturation occurring at $\omega t=\alpha$, the load voltage E_L is[1]

$$E_L=\frac{1+\cos\alpha}{2}E \tag{4}$$

and the load current I_L

$$I_L=\frac{E_L}{R_L} \tag{5}$$

For $I_X \ll I_L$, $\mathbf{i}_L$ can be approximated by (equations 1, 2, 4, and 5)

$$\mathbf{i}_L=\frac{1+\cos\alpha}{2} \tag{6}$$

as indicated in Fig. 2 by the separate ordinate axis. Since equation 6 becomes inaccurate for small load currents, the separate ordinate axis is not continued beyond $\alpha>150$ degrees.

Nomenclature

AT_A = ampere-turns acting on core A
$\mathbf{a}$ = per-unit control ampere-turns
$\Delta\mathbf{a}$ = left shift of control characteristic, per unit
B_s = saturation flux density, gausses
E_{1-N}, E_{2-N} = supply voltage, volts, average
E_L = load voltage, volts, average
ΔE = change of supply voltage, volts, average
E_m = crest (maximum) value of supply voltage, volts
e_{1-N}, e_{2-N} = supply voltage, volts, instantaneous
$e_{G,A}$, $e_{G,B}$ = gate voltages, volts, instantaneous
I_C = control current, amperes, average
i_C = control current, amperes, instantaneous
$I_{G,A}$, $I_{G,B}$ = gate currents, amperes, average
$I_{G,m}$ = maximum gate current, amperes, average
$i_{G,A}$, $i_{G,B}$ = gate currents, amperes, instantaneous
$i_{G,A,\pi}$ = gate current at $\omega t=\pi-0$, amperes, instantaneous
I_L = load current, amperes, average
$I_{L,m}$ = maximum load current, amperes, average
I_L = load current, amperes, instantaneous
$\mathbf{i}_L$ = per-unit load current, average

Paper 54-316, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 22, 1954; made available for printing April 20, 1954.

H. F. STORM is with the General Electric Company, Schenectady, N. Y.

The assistance of R. J. Noorda in obtaining the test data is gratefully acknowledged.

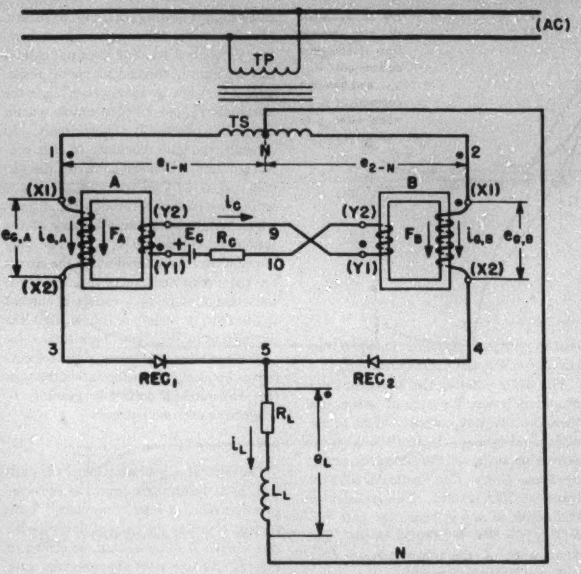


Fig. 1. Circuit diagram of magnetic amplifier with load consisting of resistance R_L and linear inductance L_L

I_S=current in shunt rectifier, amperes, average
I_S=current in shunt rectifier, amperes, instantaneous
$\mathfrak{i}_S$=per-unit current in shunt rectifier, average
L_L=load inductance, henrys
L_C=inductance of one control winding, henrys
N_C=number of turns of one control winding
N_G=number of turns of one gate winding
R_C=resistance of control circuit, ohms
R_G'=resistance of control circuit referred to gate winding, ohms
$R_{G,A}$, $R_{G,B}$=gate resistance, ohms
$R_{R,f}$=rectifier forward resistance, ohms
R_L=load resistance, ohms
X_G=gate reactance with unsaturated core, ohms
α=saturation angle, radians or degrees
φ=phase angle (equation 9), radians or degrees
ϕ_A, ϕ_B=core fluxes, maxwells, instantaneous
Φ_s=saturation flux, maxwells, instantaneous
ω=angular supply frequency, radians per second

Per-unit quantities are indicated in text by an Old English letter, and in the illustrations by a bar underneath the symbol.

Effect of Load Inductances on Control Characteristic

When the load becomes inductive, as expressed by finite ratios of $\omega L_L/R_L$ (Fig. 2), the control characteristic changes. This change is characterized by two features: 1. the control characteristic becomes shifted to the left; and 2. instability, also called snap-action, occurs.

For instance, starting at $\mathfrak{a}=0$, and increasing the control voltage in a negative direction, the control characteristic for $\omega L_L/R_L=0.105$ is traversed until point *A* is reached, where the stable region of the control characteristic terminates. By making the control voltage ever so slightly more negative, point *B* is reached and then operation continues along the so-called negative control branch *MBC*. It becomes apparent that an infinitely small change of control voltage is sufficient to produce the finite change of load current between *A* and *B*. This multivaluedness of load current for the same control signal is called instability or snap action. While the control characteristic gives the illusion of an abrupt jump between *A* and *B*, it should not be concluded that the change of load current from *A* to *B* occurs in zero time. Oscillograms indicate that it may take quite a few cycles before *B* is reached; these oscillograms show furthermore that in reality the load and control currents follow an unstable branch of the control characteristic, beginning at *A* and bending toward the minimum point *M*, as shown by the dashed curve. Somewhere near *M* the operation transfers from the unstable characteristic *AM* to the stable, negative characteristic *MBC* and the currents follow the negative characteristic until *B* is reached.

Conversely, by increasing the control voltage in a positive direction, one moves from *C* along the negative characteristic which terminates at *M*. By increasing the control voltage beyond *M* by an infinitesimal, positive amount, the load current increases from *M* to *D*, exhibiting another illusionary jump in the control characteristic. Actually, the unstable part *AM* and the stable part *AD* of the control characteristic are now traversed in the opposite direction until point *D* is reached. Again, the transfer form *M* to *D* requires a finite time.

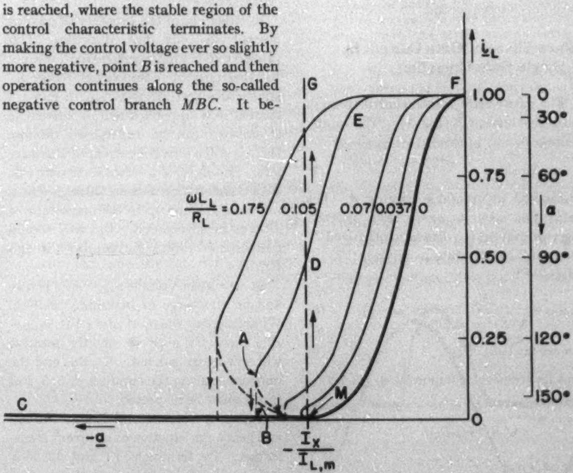


Fig. 2. Measured control characteristics obtained from circuit shown in Fig. 1

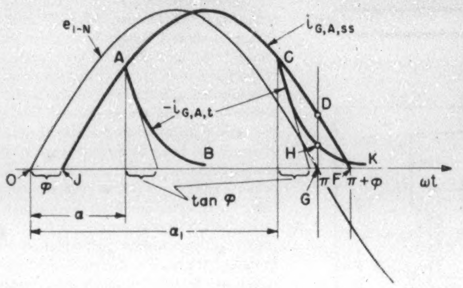


Fig. 3. Synthesis of gate current from steady-state component $i_{G,A,ss}$, and transient component $i_{G,A,t}$ when other gate circuit is open

It is concluded that the control characteristic has the stable parts *ADEF* and *CBM*, and an unstable part *AM*, the latter causing the pseudocharacteristics, *AB* and *MD;* transfer from *A* to *B* occurs only in a downward direction, and transfer from *M* to *D* only in an upward direction. The remaining characteristics of Fig. 2 are taken for various ratios of $\omega L_L/R_L$ and follow the same general pattern. It is interesting to observe that the upward jump in all characteristics occurs along the same vertical *MG*.

The saturable reactors used in these tests consist of two tape-wound cores, each core with an inside diameter of 1 inch, an outside diameter of 1.5 inch, height 0.375 inch, Orthonol, 0.004 inch tape; $N_G=2{,}600$ turns, 60 ohms, common control winding $N_C=2{,}000$ turns, 397 ohms; the rectifiers are of the General Electric germanium diffusion type *1N92;* $E=50$ volts average, $I_{L,m}=0.21$ amperes, $f=60$ cycles per second.

Wave Shape of Gate Current in Single Gate Operation

The gate circuit is interrupted through saturable reactor *B* (Fig. 1). The supply voltage

$$e_{1-N}=E_m \sin \omega t \qquad (7)$$

is applied to saturable reactor *A*. Let saturation occur at $\omega t=\alpha$ (Fig. 3), then the gate current $i_{G,A}$ can be synthesized[2-4] from its steady-state component $i_{G,A,ss}$ (*JACDF*)

$$i_{G,A,ss}=\frac{E_m}{\sqrt{R_L^2+(\omega L_L)^2}}\sin(\omega t-\varphi) \qquad (8)$$

$$\varphi=\text{arc tan}\,(\omega L_L/R_L) \qquad (9)$$

and its transient component $i_{G,A,t}(AB)$

$$i_{GA,t}=-\frac{E_m}{\sqrt{R_L^2+(\omega L_L)^2}}\epsilon^{-\frac{\omega t-\alpha}{\tan\varphi}}\sin(\alpha-\varphi) \qquad (10)$$

for $\omega t>\alpha$.

The gate current $i_{G,A}$

$$i_{G,A}=i_{G,A,ss}+i_{G,A,t} \qquad (11)$$

is the vertical distance between sine curve *JACDF* and exponential *AB*.

For $\alpha<\pi-3\tan\varphi$, the transient component is practically zero at $\omega t=\pi$, and therefore the gate current *DG* at $\omega t=\pi$, designated by $i_{G,A,\pi}$, is unaffected by the saturation angle α. For saturation angles $\pi-3\tan\varphi<\alpha<\pi$ (for instance α_1) the transient *CHK* results. This transient is still active at $\omega t=\pi$, hence the gate current $i_{G,A,\pi}=DH$, is affected by α. The significance of the magnitude of $i_{G,A,\pi}$ will soon become apparent.

Effect of Second Gate at $\omega t=\pi+0$

The gate circuit of saturable reactor *B* is re-established at $\omega t=\pi$, as shown in Fig. 1. As shown by *DG* in Fig. 3, there is a current $i_{G,A,\pi}$ flowing through the *A* gate which also flows through rectifier REC_1.

Rectifiers have the property of passing current in only one direction. This property refers to the net current, but not to the current components. For instance, if a rectifier is passing a forward current of 3 amperes a reverse current of 2 amperes can be superposed because there is still a forward current of 1 ampere left. Similarly, if a forward current *DG*, Figs. 3 and 4(A), is flowing through REC_1, a reverse current up to the magnitude of *DG* can be superposed. But why should a superposed reverse current flow through REC_1?

At $\omega t=\pi$ the voltage $e_{2-N}-e_{1-N}$ is zero and on the verge of becoming positive. What happens when, at $\omega t=\pi+0$, $e_{2-N}-e_{1-N}$ becomes ever so slightly positive will now be considered. To this end the impedance along the circuit 2, 4, 5, 3, 1 of Fig. 1 must be inspected.

The saturable reactor *B* is unsaturated, and hence can function as a current transformer. Its terminals *Y*1 and *Y*2 look into a circuit consisting of the control winding of saturable reactor *A*. Because $i_{G,A,\pi}$ is flowing (*DG*, Fig. 3) saturable reactor *A* is still saturated, and therefore its control winding does not exhibit any reactance; the control circuit resistance is very low by assumption, hence the terminals *Y*1 and *Y*2 of saturable reactor *B* look into practically a short circuit. As a result, the gate reactance of *B* is considered zero. Continuing along the circuit 2, 4, 5, 3, 1 on which $e_{2-N}-e_{1-N}$ is acting, we now pass through the *A* gate, which has zero reactance because of the saturation caused by $i_{G,A,\pi}$. The only impedances encountered along the circuit are the circuit resistances through which the voltage $e_{2-N}-e_{1-N}$ causes a current $i_{G,B}$ to flow, provided, of course, that the net current in REC_1 is still positive.

The current $i_{G,B}$ requires a control current i_C because of transformation through the unsaturated saturable reactor *B*. Neglecting exciting currents

$$i_C N_C=-i_{G,B}N_G \qquad (12)$$

The current $i_{G,B}$ flowing along the path 2, 4, 5, 3, 1 subtracts from the previous gate current. Under "previous" gate current is meant the current flowing with gate circuit *B* disconnected, as shown in Fig. 3. At $\omega t=\pi+0$ the previous gate current is smaller than $i_{G,A,\pi}$ by the infinitesimal amount by which $i_{G,A}$ has changed from $\omega t=\pi$ to $\omega t=\pi+0$. Since the following is concerned with magnitudes, but not with rates, the previous gate may be considered as constant in the vicinity of $\omega t=\pi$.

With $i_{G,B}$ flowing in opposition to the previous gate current $i_{G,A,\pi}$, the ampere-turns on core *A* decline. A further reduction of ampere-turns on core *A* occurs through the control winding (equation 12). The ampere-turns acting on core *A* are designated by AT_A

$$AT_A=i_{G,A,\pi}N_G-i_{G,B}N_G+i_C N_C \qquad (13)$$

Substituting from equation 12

$$AT_A=(i_{G,A,\pi}-2i_{G,B})N_G \qquad (14)$$

Letting $\omega t>\pi+0$, the current $i_{G,B}$ will increase in proportion to $e_{2-N}-e_{1-N}$, and AT_A will decrease. For

$$i_{G,B}=\frac{1}{2}i_{G,A,\pi} \qquad (15)$$

$AT_A=0$, indicating that core *A* becomes unsaturated. It will be noted that at this instant the net forward current $i_{G,A}$ in REC_1 is still positive, namely

$$i_{G,A}=i_{G,A,\pi}-i_{G,B}=\frac{1}{2}i_{G,A,\pi} \qquad (16)$$

and hence the superposition of $i_{G,B}$ in the reverse direction of REC_1 was proper.

In the absence of reactances, the

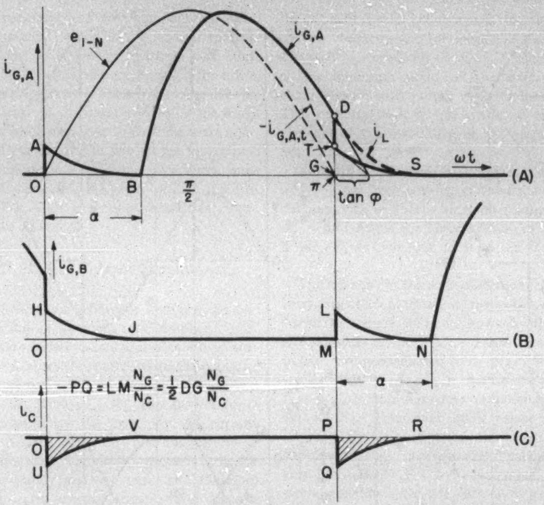


Fig. 4. Gate currents $i_{G,A}$, $i_{G,B}$, and extra control current i_C, when both gate circuits are connected

amount of time it takes from $\omega t=\pi$ until core A unsaturates (equation 15) depends on $e_{2-N}-e_{1-N}$ and the circuit resistance. In order to avoid overcomplications, it will be assumed that the resistance of circuit 2, 4, 5, 3, 1, 2 is zero. Then, it is concluded that at $\omega t=\pi+0$:

1. Both cores are unsaturated.
2. The load current $i_{G,A,\pi}$ splits into two equal components (equations 15 and 16).
3. The gate currents $i_{G,A}$ and $i_{G,B}$ do not encounter reactance by flowing from N to 5 (Fig. 1), because control ampere-turns equal but opposite to the gate ampere-turns (equation 12) are flowing.

Exciting Intervals ($0<\omega t<\alpha$, $\pi<\omega t<\pi+\alpha$)

The excitation process with purely resistive load has been discussed in reference 1. For instance, during the interval $\pi<\omega t<\pi+\alpha$ core A becomes excited from its control winding, which, in turn, is energized from the control winding of saturable reactor B. As a result, a voltage appears at the gate of saturable reactor A. If no transformation losses are involved, the gate voltage of A is identical with the gate voltage of B, and hence there is no voltage between N and 5 of Fig. 1.

If a current is now injected at N, such as coming from the load inductance L_L, this current is expected to flow to 5 without encountering reactance. The load current entering at N splits up into two components $i_{G,A}$ and $i_{G,B}$; see Fig. 4. Since both cores are unsaturated, transformation between gate and control windings takes place. The control current is common to both saturable reactors and therefore (neglecting the exciting currents)

$$i_{G,A}=i_{G,B}=\frac{1}{2}i_L \tag{17}$$

The current wave shapes of Fig. 3 were obtained with the gate circuit of B open. The supply voltage e_{1-N} was active as long as the gate current $i_{G,A}$ was flowing (until $\omega t=\pi+\varphi$). With circuit B closed, it was found, however, that there is no voltage between N and 5, and hence no voltage across the load ($e_L=0$) during the exciting intervals. Therefore, the load current i_L will not continue as a part of a sinusoid, but will decline exponentially with the time constant $\omega L_L/R_L=\tan\varphi$ as shown by equations 9 and 10 ($\alpha=\pi$).

This is demonstrated in Fig. 4(A). For $\alpha<\omega t<\pi$, the load current i_L is identical with the gate current $i_{G,A}(DG)$. For $\pi<\omega t<\pi+\alpha$ the load voltage $e_L=0$, and therefore i_L begins to decline exponentially (dotted curve DS). From equations 17 $i_{G,A}$ [TS, Fig. 4(A)] and $i_{G,B}$ [LN, Fig. 4(B)] are found and it is seen that both gate currents flow simultaneously. It should be noted, however, that one half of the load current flows in the opposite direction of the transformer secondary voltage. For instance, during $\pi<\omega t<\pi+\alpha$, $i_{G,A}$ is opposed by e_{1-N}, but $i_{G,A}$ can flow because of the gate voltage A obtained by transformation from the B gate: With no losses present, the gate voltage A is equal in magnitude, but opposite in sign of the transformer voltage e_{1-N}.

Because of the symmetry of the circuit similar conditions will also be found in the interval from $0<\omega t<\alpha$, and corresponding gate currents AB in Fig. 4(A) and HJ in Fig. 4(B). Exciting currents are not shown in Fig. 4.

In reality the circuit resistances are not zero, and therefore the current fronts at $\omega t=0$ and π are not vertical, but exhibit finite slopes.

An oscillogram of supply voltage and gate current is shown in Fig. 5.

CONCLUSIONS FOR EXCITING INTERVALS

1. The conclusions for $\omega t=\pi+0$ also hold throughout the exciting intervals.
2. Despite the fact that the gates carry components of the load current the mechanism of one core being excited by the other core is the same as for purely resistive load.[1] It follows that the core fluxes and the exciting components of the control current are also the same as for purely resistive load.
3. The load voltage e_L (N to 5, Fig. 1) is zero during the exciting intervals, as in the case with purely resistive load.

Saturation Intervals ($\alpha<\omega t<\pi$, $\pi+\alpha<\omega t<2\pi$)

Let core B saturate at $\omega t=\pi+\alpha$. There is no longer any voltage transformed from the B gate into the A gate, via control circuit. Without the transformed voltage, the gate current $i_{G,A}$

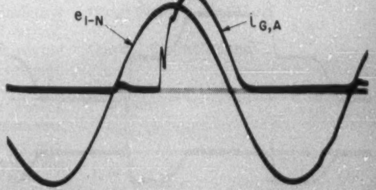


Fig. 5. Oscillogram. Supply voltage e_{1-N}, gate current $i_{G,A}$

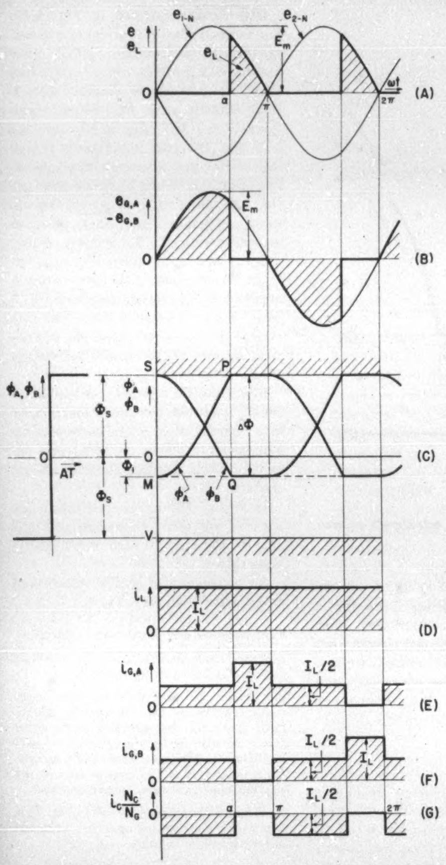

e
e_L
e_1-N
e_2-N
e_L
E_m
ωt
0
a
π
2π
(A)
e_G,A
-e_G,B
E_m
0
(B)
φ_A, φ_B
S
P
φ_A
φ_B
Φ_S
ΔΦ
O
0
Φ_1
AT
M
Q
φ_A
φ_B
Φ_S
V
(C)
i_L
I_L
0
(D)
i_G,A
I_L
I_L/2
0
(E)
I_L/2
i_G,B
I_L
0
(F)
i_C N_C/N_G
0
I_L/2
a
π
2π
(G)

Fig. 6. Wave shapes occurring in circuit of Fig. 1, when $\omega L_L \gg R_L$. Exciting currents are not shown. Cores saturate with positive polarity

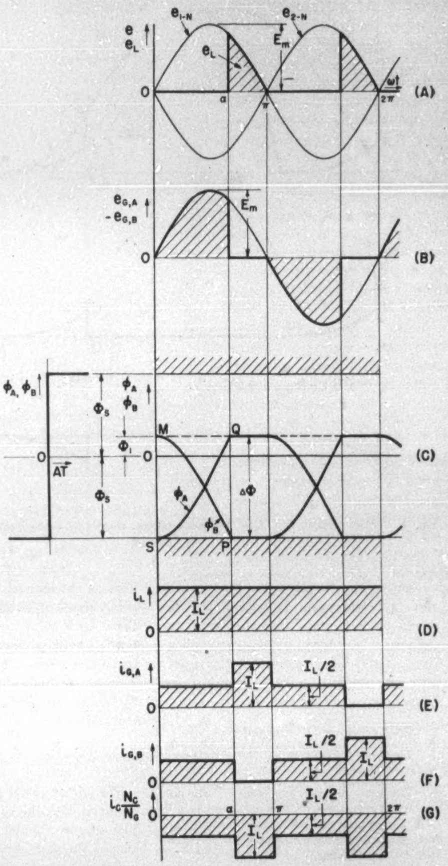

e
e_L
e_1-N
e_2-N
e_L
E_m
ωt
0
a
π
2π
(A)
e_G,A
-e_G,B
E_m
0
(B)
φ_A, φ_B
φ_A
φ_B
Φ_S
M
Q
Φ_1
O
0
AT
φ_A
Φ_S
ΔΦ
φ_B
S
P
(C)
i_L
I_L
0
(D)
i_G,A
I_L
I_L/2
0
(E)
I_L/2
i_G,B
I_L
0
(F)
i_C N_C/N_G
0
I_L/2
a
π
2π
I_L
(G)

Fig. 8. Wave shapes occurring in circuit of Fig. 1, when $\omega L_L \gg R$. Exciting currents are not shown. Cores saturate with negative polarity

cannot flow in opposition of the supply voltage. Therefore gate current $i_{G,A}$ stops, and gate current $i_{G,B}$ becomes equal to the load current. The load voltage e_L is the line-to-neutral voltage of the transformer secondary *TS* (Fig. 1).

CONCLUSIONS FOR SATURATION INTERVALS

1. The load voltage E_L for inductive load is identical with E_L for purely resistive load and hence equation 4 is also valid for inductive d-c load.
2. It follows from conclusion 1 that equations 5 and 6 are also valid for inductive d-c load.

Extra Control Current

The total control current of the amplifier with inductive d-c load can be viewed

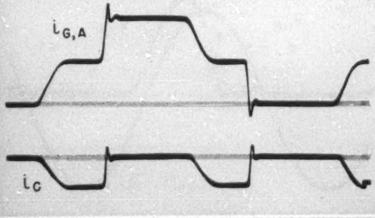
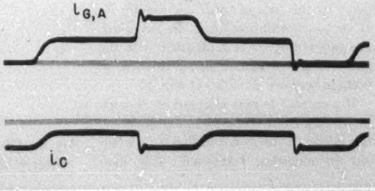


Fig. 7 (left). Oscillogram. Gate current $i_{G,A}$, control current i_C

Fig. 9 (right). Oscillogram. Gate current $i_{G,A}$, control current i_C



as consisting of two components: one component which is the normal control current with purely resistive load, and an additional component, called the extra control current, which owes its existence entirely to the inductiveness of the load. It will be seen shortly that the extra control current is the key to the instability of magnetic amplifiers. The extra control current i_C is shown by UV and QR in Fig. 4(C) and is obtained from equations 12 and 17.

Left Shift of Control Characteristic

The major difference between the operation of the magnetic amplifier with purely resistive load and with a load also containing inductance lies in the presence of the extra control current. Depending on the ratio $\omega L_L/R_L$, the ampere-turns produced by the extra control current are of the same order of magnitude or larger than the control ampere-turns for the purely resistive case, and therefore the extra control ampere-turns have a crucial bearing on the deviation of the control characteristic with inductive load from the control characteristic with purely resistive load. The extra control current is always negative and therefore the control characteristic with inductive load becomes shifted to the left of the control characteristic with purely resistive load.

To obtain a better quantitative insight, the analysis is now subdivided into loads with small time constants and loads with long time constants.

$\omega L_L \ll R_L$; See Fig. 2

CAUSE OF INSTABILITY

As shown in Fig. 3, the saturation angle α can be varied considerably without causing the gate current $i_{G,A,\pi}$ to change. As long as $i_{G,A,\pi}$ [DG, Figs. 3 and 4(A)] remains constant, the extra control current remains constant, and hence the left shift with respect to the case with purely resistive load ($\omega L_L=0$) will also remain constant.

As the saturation angle increases over $\pi-3\tan\varphi$ the starting transient $i_{G,A,t}$ will cause the gate current $i_{G,A,\pi}$ (for instance, HD in Fig. 3) to decline. As a result, the absolute value of the extra control current also becomes smaller. It follows that the left shift of the control characteristic becomes smaller when α increases over $\pi-3\tan\varphi$. As a result, the control characteristic curves back toward M (dashed curves in Fig. 2). For $\alpha=\pi$, the extra control current is zero, and hence the left shift is zero, which means that the control characteristic for inductive load curved back completely to point M, which is the minimum point of the control characteristic with purely resistive load. This curving back is the graphical expression of the instability of the amplifier.

Since the phase angle φ increases with $\omega L_L/R_L$, the back-curving of the control characteristic will occur for smaller values of α, and hence for higher values of i_L if $\omega L_L/R_L$ is increased, as shown in Fig. 2.

The left shift of the control characteristic is proportional to the average value of the extra control current, shown by the shaded areas in Fig. 4(C). These areas, in turn, are proportional to $I_{G,A,\pi}{}^2$. For small values of $\omega L_L/R_L$ it can be shown that $i_{G,A,\pi}$ is approximately proportional to $\omega L_L/R_L$. Hence for small values of $\omega L_L/R_L$ the left shift of the control characteristic is approximately proportional to $(\omega L_L/R_L)^2$. In the presence of circuit resistance, the left shift $\Delta\alpha$, in per-unit notation, can be approximated (without proof) by

$$\Delta\alpha \doteq \frac{1}{2}\left(\frac{\omega L_L}{R_L}\right)^2 \frac{1}{1+\dfrac{\frac{1}{2}(R_G+R_{R,f})+\frac{1}{4}R_C'}{R_L}} \tag{18}$$

for $\omega L_L \ll R_L$.

INSTABILITY WITH PURELY RESISTIVE LOAD

By connecting small inductances into the gate circuits, the control characteristic will display instability, even if the load is purely resistive ($L_L=0$).

For small values of $\omega L_L/R_L$ and for values of α in the vicinity of π, it makes little difference whether the inductance L_L is located in the load circuit, as shown in Fig. 1, or if the inductance L_L is inserted in each gate lead (3, 4) instead. It follows that a magnetic amplifier with purely resistive load may also display instability because of saturation and leakage reactance of the gate windings. (Snap action with purely resistive load may also occur because of the shape of the dynamic hysteresis loop.)

$\omega L_L \gg R_L$

As $\omega L_L/R_L$ increases, the analytical expression for the extra control current, and hence the prediction of the control characteristic, becomes more complex. However, if $\omega L_L/R_L$ becomes large, for instance, 30 radians or more, the problem again becomes quite simple, because then the load current can be considered continuous and free of ripple, as shown in Fig. 6(D).

Exciting Intervals ($0<\omega t<\alpha$, $\pi<\omega t<\pi+\alpha$)

From equations 12 and 17 the gate currents and the extra control current can be obtained immediately; see Fig. 6(E), (F), (G). The extra control current with inductive load is so large with respect to the control current with purely resistive load that the latter is neglected. Again $e_{G,A}=-e_{G,B}$ and $e_L=0$ see Fig. 6(A), and (B). The core fluxes Fig. 6(C), are the same as for purely resistive load.[1]

Saturation Intervals ($\alpha<\omega t<\pi$, $\pi+\alpha<\omega t<2\pi$)

CORE FLUXES SATURATE WITH POSITIVE POLARITY

At $\omega t=\alpha$ saturable reactor A saturates [P, Fig. 6(C)] and there is no longer any voltage transformed from the A gate into the control circuit, and hence into the B gate. Therefore the gate current $i_{G,B}$ stops, and gate current $i_{G,A}$ becomes equal to the load current i_L. With saturable reactor A saturated and no gate current flowing in the unsaturated saturable reactor B, there can be no gate current transformed into the control circuit and the control current $i_C=0$. Regardless of the ratio L_L/R_L the load voltage E_L is again given by equation 4 and the load current by equations 5 and 6.

Also, as previously pointed out, the core fluxes ϕ_A and ϕ_B Fig. 6(C), are the same as for purely resistive load[1] and therefore they are independent of the ratio L_L/R_L.

An oscillogram of gate and control current is shown in Fig. 7.

CORE FLUXES SATURATE WITH NEGATIVE POLARITY

A different mode of operation results if the core fluxes are saturating with negative polarity, as shown in Fig. 8(C). The wave shapes shown in (A), (B), (D), (E), and (F) or Fig. 8 are identical with the similarly designated wave shapes of Fig. 6. The only wave exhibiting a different shape is the control current; see Figs. 6(G), and 8(G).

An oscillogram of gate and control current is shown in Fig. 9.

Control Characteristic

CORE FLUXES SATURATE WITH POSITIVE POLARITY

From Fig. 6(G)

$$I_C=-\frac{1}{2}I_L\frac{N_G}{N_C}\frac{\alpha}{\pi} \tag{19}$$

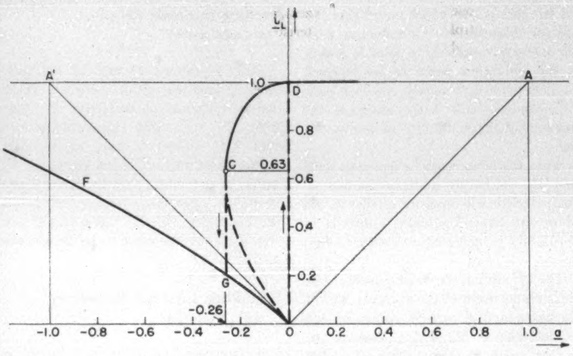


Fig. 10. Normalized characteristic for $\omega L_L \gg R_L$

From equations 1, 3, and 5

$$\mathbf{a} = -\mathbf{i}_L \frac{\alpha}{\pi} \tag{20}$$

Substituting for α from equation 6, the control characteristics are obtained in normalized form

$$\mathbf{a} = -\mathbf{i}_L \left[\frac{1}{\pi} \operatorname{arc\,cos}\,(2\mathbf{i}_L - 1) \right] \tag{21}$$

This control characteristic is shown by *DCG* in Fig. 10.

CORE FLUXES SATURATE WITH NEGATIVE POLARITY

In analogy to equations 20 and 21

$$\mathbf{a} = -\mathbf{i}_L \left(2 - \frac{\alpha}{\pi} \right) \tag{22}$$

$$\mathbf{a} = -\mathbf{i}_L \left[2 - \frac{1}{\pi} \operatorname{arc\,cos}\,(2\mathbf{i}_L - 1) \right] \tag{23}$$

This portion of the control characteristic is shown by *OGF* in Fig. 10.

COMBINED CHARACTERISTICS

The combined control characteristic of the circuit of Fig. 1 is *FGOCD*. It can be recognized immediately that the characteristic is again multivalued. (For the purpose of comparison, the control characteristic of the ordinary, parallel-connected saturable reactor with purely resistive load is shown by the straight lines *OA* and *OA'*, extending angles at 45 degrees with the abscissa axes.) The unstable branch of the characteristic is *CO*. The co-ordinates *C* of one end point of the unstable branch are $\mathbf{a} = -0.26$ and $\mathbf{i}_L = 0.63$, as shown in the Appendix. The other end point is *O* when exciting currents are neglected; otherwise it is *M*, as shown in Fig. 2.

Actually measured control characteristics are shown in Fig. 11 for various ratios $\omega L_L/R_L$.

Elimination of Snap Action

There are cases where snap action of a magnetic amplifier is desirable, i.e., to obtain control properties similar to a latch-in type relay. The preponderant case, however, is where snap action is not acceptable, i.e., in processes which require gradual control by hand. In closed-loop control systems, snap action does not necessarily lead to system instability, but snap action certainly increases antihunt problems. No wonder, then, that a great deal of attention has been given to the elimination of snap action, and several methods for the elimination of snap action are discussed in the following.

It has been shown in the preceding part that snap action occurs when load current is transformed from the gate into the control windings during the exciting interval. Means to combat snap action must therefore be directed to prevent this transformation, or render it ineffective, or to prevent the amplifier from operating within the zone of instability.

Operation of Amplifier Outside of Zone of Instability

Instability has been previously traced to the reduction of the gate current $i_{G.A,\pi}$ with increasing saturation angles. By preventing $i_{G.A,\pi}$ from dropping below

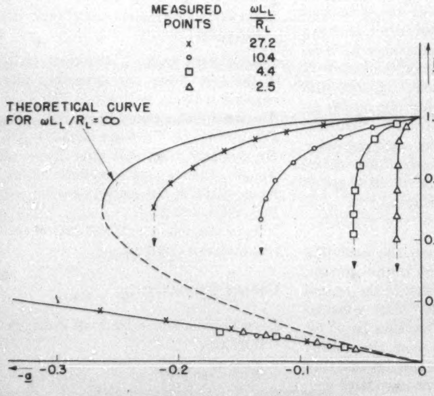


Fig. 11. Measured control characteristics obtained from circuit of Fig. 1. Instability in upward direction as in Fig. 2, MG

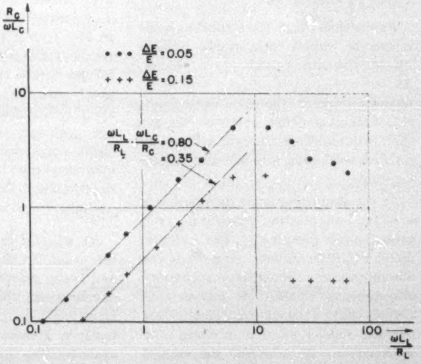


Fig. 12. Minimum values for $R_C/\omega L_C$ to prevent instability of circuit shown in Fig. 1

a certain minimum value, instability is averted. This method is explained in the following.

The supply voltage E (indices omitted) has been assumed to swing the flux density in the cores exactly from minus saturation to plus saturation (or vice versa). This is another way of saying that the maximum reactive voltage (counterelectromotive force) the gates can develop is E. Now, if the supply voltage is raised by ΔE volts, this voltage ΔE must spill over into the load circuit, causing the flow of an irreducible minimum load current $I_{L,\min}$

$$I_{L,\min} = \frac{\Delta E}{R_L} \tag{30}$$

Therefore, by the simple expedient of raising the supply voltage, a minimum value for $i_{G,A,\pi}$ can be maintained regardless of the magnitude of the control current, and hence instability can be avoided. This procedure is most practical for comparatively small values of $\omega L_L/R_L$.

Impeding Extra Control Current

By introducing more resistance into the control circuit, and preferably also increasing the supply voltage, snap action can be avoided. Fig. 12 shows how the minimum control circuit resistance R_C for snap-free operation is related to $\omega L_L/R_L$ and the ratio $\Delta E/E$; L_C is the inductance of the control winding based on one unsaturated core. Over a considerable range the product $(\omega L_L/R_L)$ $(\omega L_C/R_C)$ is constant.

The introduction of more resistance R_C into the control circuit reduces the power gain. Where this reduction is to be avoided, an external inductance can be inserted in the control circuit instead; however, a drawback of adding an inductance may be seen in the increase of the control time constant.

Preventing Generation of Extra Control Current

Suppose a synchronous switch would short $5-N$, Fig. 1, during the exciting intervals. The inductance L_L would maintain the load current through the short circuit and therefore the gate currents would stop flowing when the supply voltage is zero (provided there is no saturation or air-core reactance). As a result no extra control current flows and instability is averted. The excitation of the cores is the same as for a purely resistive load.[1] Hence the relation between saturation angle α and control ampere-turns $I_C N_C$ is the same as for purely resistive load. The load voltage (equation 4) is the same in both cases, and so is the load current I_L (equations 5 and 6). One concludes that the magnetic amplifier with inductive load, Fig. 1, and with a synchronous short circuit during the exciting interval has the same control characteristic as the magnetic amplifier with purely resistive load and without synchronous switch.

Magnetic Amplifier with Shunt Rectifier[5]

It will be seen that the load current flows always in the same direction, namely from N to 5 in the previously mentioned short circuit. Therefore the synchronous switch can be replaced by the rectifier REC_3, as shown in Fig. 13. This rectifier is called: free-wheeling rectifier, by-pass rectifier, shunt rectifier. Provided the forward resistance of the shunt rectifier REC_3 is low, the magnetic amplifier with REC_3 will function as the magnetic amplifier with the synchronous contact. The control characteristic of the latter is practically identical with that of the magnetic amplifier with purely resistive load, hence the control characteristic of the magnetic amplifier with inductive load and with shunt rectifier, Fig. 13, is practically identical with the control characteristic of the magnetic amplifier with purely resistive load, provided the forward resistance of shunt rectifier REC_3 is very low. If the forward resistance of REC_3 becomes more pronounced, a voltage appears between 5 and M, leading to a left shift of the control characteristic, Fig. 2, and to snap action.

For zero rectifier forward resistance and a very large ratio of $\omega L_L/R_L$, the wave shapes of Fig. 14 result. The average value I_S of the shunt rectifier current follows from Fig. 14(G):

The per-unit shunt current is designated by $\mathbf{i}_S$

$$I_S = I_L \frac{\alpha}{\pi} \tag{31}$$

$$\mathbf{i}_S = \frac{I_S}{I_{L,m}} \tag{32}$$

Then

$$\mathbf{i}_S = \mathbf{i}_L \frac{\alpha}{\pi} \tag{33}$$

and from equation 20

$$\mathbf{i}_S = |\mathbf{a}| \tag{34}$$

Therefore equation 21 applies also to $\mathbf{i}_S$, and the relation $\mathbf{i}_S$ versus $\mathbf{i}_L$ in Fig. 15 is obtained by replotting curve OCD of Fig. 10. Hence, the maximum current

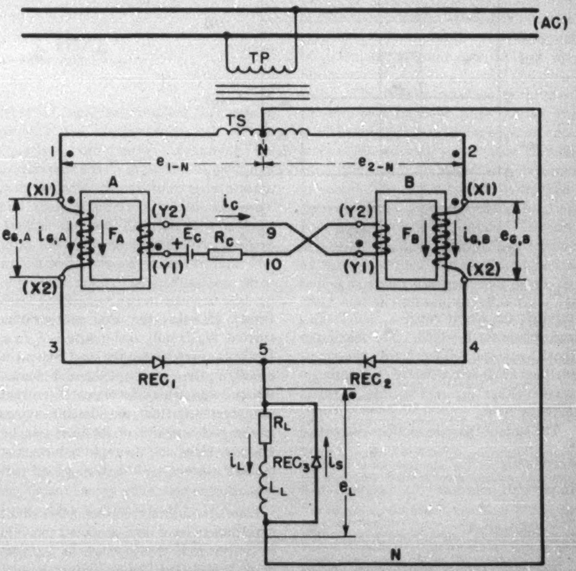


Fig. 13. Circuit diagram of magnetic amplifier similar to Fig. 1, but with the addition of REC_3, called free-wheel, by-pass, shunt rectifier

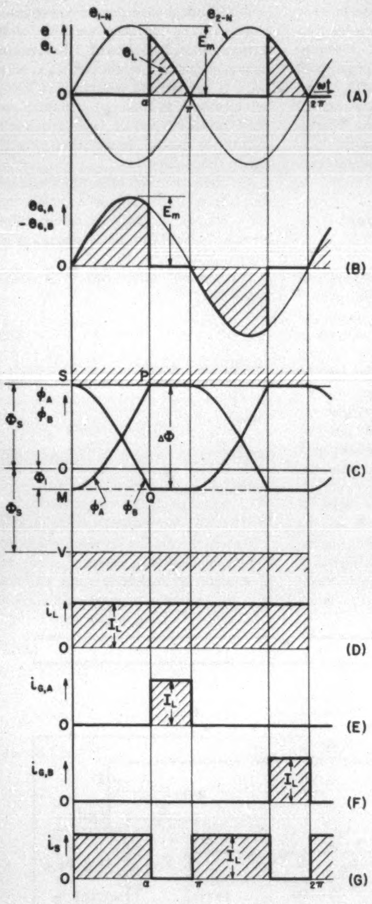


Fig. 14 (left). Wave shapes occurring in circuit of Fig. 13, when $\omega L_L \gg R_L$. Exciting currents are not shown

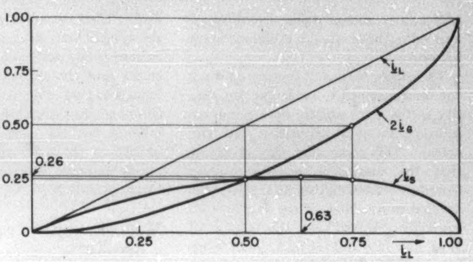


Fig. 15 (above). Per-unit shunt rectifier current $\mathbf{i}_S$, and per-unit gate rectifier currents $\mathbf{i}_L - \mathbf{i}_S = 2\mathbf{i}_G$, as function of per-unit load current $\mathbf{i}_L$

through the shunt rectifier is 0.26 $I_{L,m}$ and occurs for $\mathbf{i}_L=0.63$. The maximum (peak) reverse voltage across the shunt rectifier REC_3 is one-half of maximum reverse voltage across the gate rectifiers REC_1 or REC_2.

The sum of the gate rectifier currents

$$2I_G = I_L - I_S \tag{35}$$

In per-unit notation

$$\frac{2I_G}{I_{L,m}} = 2\mathbf{i}_G = \mathbf{i}_L - \mathbf{i}_S \tag{36}$$

The gate rectifier currents are always smaller than the load current (except for $\mathbf{i}_L=1$). For instance, at the mid-characteristic ($\mathbf{i}_L=0.5$) the total gate rectifier current $2I_G$ is only one-quarter of $I_{L,m}$. In applications where the load current is usually in the vicinity of the mid-characteristic, and where increases of current are only intermittent, considerable savings in the cost and size of rectifiers can be made by observing the rapid reduction of rectifier current with declining load current.

Instead of a shunt rectifier, other shunt admittances have been proposed to avoid or reduce snap action, such as a resistance, a capacitor, or an inductance-capacitance filter tuned to 2 *f*, connected in parallel to the load inductance.

Appendix

To obtain the extreme point *C*, Fig. 10

$$\frac{d\mathbf{i}}{d\mathbf{i}_L} = \frac{d\mathbf{i}}{d\alpha}\frac{d\alpha}{d\mathbf{i}_L} = 0 \tag{24}$$

Differentiating equations 20 and 6

$$\frac{d\mathbf{i}}{d\alpha} = \alpha \sin \alpha - 1 - \cos \alpha = 0 \tag{25}$$

$$\alpha = \cot \frac{\alpha}{2} \tag{26}$$

$$\alpha \doteq \frac{5}{12}\pi \text{ radians} \doteq 75 \text{ degrees} \tag{27}$$

From equation 6

$$\mathbf{i}_L = \frac{1+\cos 75 \text{ degrees}}{2} = 0.63 \tag{28}$$

From equation 20

$$\mathbf{i} = -0.63\frac{5}{12} = -0.26 \tag{29}$$

References

1. Theory of Magnetic Amplifiers with Square-Loop Core Materials, H. F. Storm. *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 629–40.

2. Electric Circuits (book), M.I.T. Electrical Engineering Staff. John Wiley and Sons, Inc., New York, N. Y., 1940, p. 341.

3. The Operation of Magnetic Amplifiers with Various Types of Load. Part I—Load Currents for Given Angle of Firing, L. A. Finzi, R. R. Jackson. *AIEE Transactions*, vol. 73, pt. I, July 1954, pp. 270–78.

4. The Operation of Magnetic Amplifiers with Various Types of Loads. Part II—Controlling the Angle of Firing. The Transfer Characteristics of Amplifiers with Low-Control Impedance. *Ibid.*, pp. 279–88.

5. The Transductor Amplifier (book), U. H. Krabbe. Lindhska Boktryckeriet, Orebro, Sweden 1947.

◆

Discussion

L. A. Finzi (Carnegie Institute of Technology, Pittsburgh, Pa.): Dr. Storm describes quite clearly the mechanism of instabilities caused by inductive loads in center-tap (d-c output) amplifiers. It is

Vo
Pt.
19

gratifying to see that his point of view agrees substantially with the treatment made by R. R. Jackson and myself in section 6 of reference 3 and section 8 of reference 4 of the paper.

For instance, the experimental curves of Fig. 2 of Dr. Storm's paper show that for moderate values of $\omega L_L/R_L$ the effect of inductive loads over regions of high output (early firing) is to translate the transfer characteristic obtained for resistive loads to the left by an essentially constant amount. This finding agrees with our analysis applied to the same range of the variables; in fact, equation 9 and Fig. 3 of reference 3 indicates that for a purely resistive-inductive load ($K=0$, i.e., $i_{L,\pi}=aV_{gm}/R$) with $\theta=\varphi=10$ degrees the load current $i_{L,\pi}=i_{G.A.\pi}$ which is found in the load at the beginning of the relaxation interval does not vary much by varying the angle of firing. Accordingly, the additional control current requirements due to the transformer action of relaxation currents through both gate windings are expected to be nearly constant for all α's, between, e.g., $\alpha=0$ degrees and $\alpha=130$ degrees.

On the other hand, as α increases beyond 130 degrees approaching 180 degrees, it is found that $i_{L,\pi}$ decreases; accordingly, the requirements of additional control currents compensating the magnetomotive force of relaxation currents become smaller. Thus the lower part of the transfer characteristic is expected to bend inwards (towards the point of minimum output at $\alpha=180$ degrees) as shown in the dashed portion of Dr. Storm's measured curves.

H. P. Gluckman (Department of Water and Power, City of Los Angeles, Los Angeles, Calif.): Dr. Storm's paper will be of great interest to engineers engaged in the application of magnetic amplifiers to the control of field windings of a-c and d-c generators. The use of center-tap magnetic amplifiers with highly inductive loads such as field windings will probably cause snap action and resulting instability unless remedial measures are taken. The remedy recommended by Dr. Storm is a rectifier shunting the inductive load to provide a by-pass path for the extra control current. However, the bridge-type magnetic amplifier inherently provides a by-pass path for the extra control current. Does the center-tap circuit with a free-wheeling rectifier offer substantial advantages over a self-saturating bridge circuit?

When reversal of output polarity is required, the free-wheeling rectifier is not applicable, and some other means of providing a path for the extra control current is required. The use of a second-harmonic filter across the inductive load has been suggested. Another suggestion is the use of a shunting resistor. Data on the relative efficacy of these methods would be of value to the application engineer.

H. F. Storm: I wish to thank Prof. Finzi for his discussion. It is very gratifying that he is in agreement with my paper.

In answer to Mr. Gluckman's question, the center-tap circuit with shunt rectifier offers advantages over the bridge type when, e.g., the theoretical, maximum output voltage $E_{0,m}$ of the magnetic amplifier is equal to, or is smaller than, $(1/\pi)$ times the maximum, inverse voltage $E_{i,m}$ of the rectifiers. In this case the bridge circuit requires four rectifier cells, whereas the center-tap circuit requires only three rectifier cells. Of these three rectifier cells, two cells have the same current rating as those for the bridge type, and one cell, namely the shunt rectifier, may have only one-half of that current rating.

An Analogue Computer for Automatic Determination of System Swing Curves

D. W. C. SHEN
MEMBER AIEE

S. LISSER
NONMEMBER AIEE

IN USING the conventional a-c network analyzer to solve transient-stability problems, it is necessary to resort to a step-by-step process which is both time-consuming and tedious although the power flow data can be obtained at each step directly from the analyzer. The subject of the rapid evaluation of system swing curves has recently attracted the engineering interest and needs the attention that it now seems to be receiving. This paper describes the design features of a 10-kc network analyzer in which computing elements are incorporated in the feedback loop which relates the electric power output with the phase shifter in such a way that a dynamic representation for a transient-stability solution can be directly obtained on the analyzer itself.

Boast and Rector[1] reported an analogue method for obtaining directly on calibrated d-c cathode-ray oscilloscopes the swing curves for power systems during disturbance conditions. Their method, however, does not use the network analyzer except for the evaluation of the initial boundary conditions. Similar to the differential analyzer method of solving directly the electromechanical system equations, information concerning the system parameters must first be measured from the network analyzer representation and then transferred to the analogue. Another method, described by Robert,[2] using model machines and model systems with a scaling-down process to take care of machine time constants has been criticized by Kusko[3] as not being a working analyzer suitable for everyday power system studies.

At the University of Adelaide work has been undertaken as a research project to build up small network analyzers for local utilities to enable them to study transient-stability problems without the necessity of the step-by-step procedure. One of the methods proposed is the independent energy-store representation recently described by Kaneff.[4] In Kaneff's scheme each machine representation consists of a reactance-tube-controlled oscillator synchronized to a master oscillator under steady-state operating conditions. During disturbances, the power difference causes the energy-store circuit either to store energy in, or absorb energy from, the store. This causes the instantaneous frequency of the machine oscillator to vary according to the square root of the stored energy. Swing curves are then observed on an oscilloscope by means of a limiting technique with fault application and switching carried out cyclically. The authors have been working on the idea of a dynamic phase shifter responsive to signals so that a network analyzer equipped with these phase shifters can represent both static and dynamic systems. The latter include both transient-stability problems and hunting phenomena in power systems.[5] Beckey and Schott[6] indicate that in their scheme an analogous idea was employed and a human operator was used as a power-to-shaft motion transducer. For the analyzer operators and designers who are

---

Paper 54-275, recommended by the AIEE Computing Devices Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21–25, 1954. Manuscript submitted March 25, 1954; made available for printing May 19, 1954.

D. W. C. SHEN is with the University of Adelaide, Adelaide, South Australia, and S. LISSER is with the Electricity Trust of South Australia, Adelaide, South Australia.

The authors are indebted to Prof. E. O. Willoughby for suggesting this project and for his constant interest and encouragement. To the Electricity Supply Association of Australia, the authors are under great obligation for the provision of funds made available through the Electrical Research Board of the C.S.I.R.O. Thanks are due the Electricity Trust of South Australia for co-operation throughout this investigation.

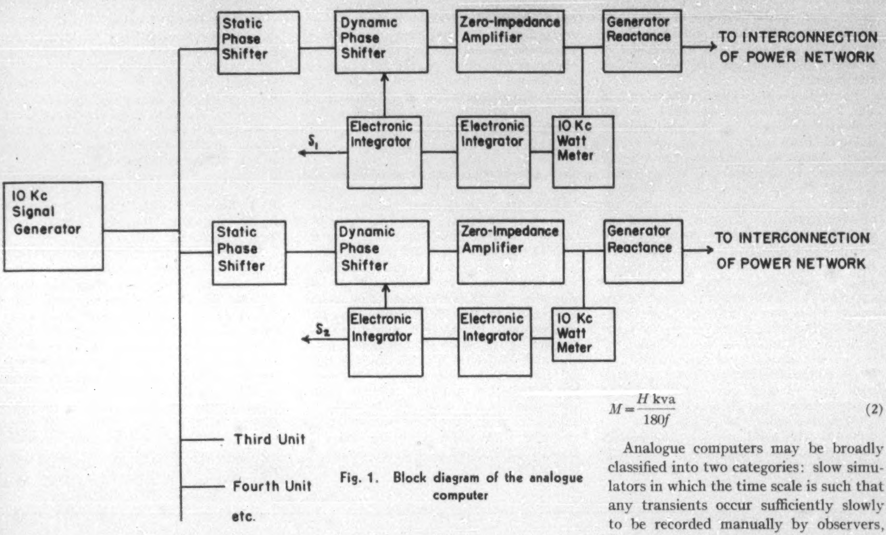

10 Kc Signal Generator
Static Phase Shifter
Dynamic Phase Shifter
Zero-Impedance Amplifier
Generator Reactance
TO INTERCONNECTION OF POWER NETWORK
Electronic Integrator
Electronic Integrator
10 Kc Watt Meter
S_1
Static Phase Shifter
Dynamic Phase Shifter
Zero-Impedance Amplifier
Generator Reactance
TO INTERCONNECTION OF POWER NETWORK
Electronic Integrator
Electronic Integrator
10 Kc Watt Meter
S_2
Third Unit
Fourth Unit
etc.

Fig. 1. Block diagram of the analogue computer

interested in the design features, this paper gives some details and also makes some suggestions for future developments with a view to achieving a higher degree of accuracy in the experimental results.

Description of Method

The principle of the present scheme is quite general; it can be applied to both the conventional 400-cycle network analyzer and the modern 10-kc analyzer described by Ryder and Boast.[7] Because of the considerably lowered cost of the high-frequency analyzer and its proven high quality of performance, the individual units in the present scheme are designed to operate at 10 kc so that well-known electronic techniques can be readily applied with advantage.

Fig. 1 shows a block diagram of the units. The reference of the system is a 10-kc crystal standard-frequency oscillator. Coupled to this oscillator are electronic generators, each consisting of a zero-impedance amplifier controlled by a dynamic phase shifter. Under steady-state conditions, the output voltages of the amplifiers are adjusted both in magnitude and in phase to correspond to the internal electromotive forces of the synchronous machines in an actual system to be represented. An electronic wattmeter is connected to the output of each amplifier to indicate the steady-state power under steady-state conditions and, with proper adjustment, to indicate accelerating power under system disturbances. The wattmeter output when indicating accelerating power is a varying direct voltage which is converted through double integration by means of electronic integrators first to represent the instantaneous angular velocity and then to represent the angular variations. These angle variations are fed back into dynamic phase shifters to swing the phase angles and to control the powers of the electronic generators. The transients which the phase shifters undergo will satisfy the following system swing equations[8]

$$M_1\frac{d^2\delta_1}{dt^2}=P_{a1}=P_{m1}-P_{e1}$$

$$M_2\frac{d^2\delta_2}{dt^2}=P_{a2}=P_{m2}-P_{e2}$$

$$\cdot$$

$$\cdot \qquad (1)$$

$$\cdot$$

$$M_n\frac{d^2\delta_n}{dt^2}=P_{an}=P_{mn}-P_{en}$$

where

M = angular momentum
δ = angular displacement between the internal voltage of the synchronous machine and an arbitrary synchronously rotating reference frame
P_a = accelerating power
P_m = mechanical power
P_e = electric power
f = normal frequency of the actual system

In terms of the inertia constant H, the angular momentum becomes

$$M=\frac{H\text{ kva}}{180f} \qquad (2)$$

Analogue computers may be broadly classified into two categories: slow simulators in which the time scale is such that any transients occur sufficiently slowly to be recorded manually by observers, and fast simulators in which transients are recorded with an oscillograph. To establish the relationship between the parameters of the computer elements and those of the actual system, let t_1 denote the time in the simulator which is related to the time t in the actual problem by a scale factor σ

$$t_1=\sigma t \qquad (3)$$

Referring to Fig. 2, ΔP, the accelerating power in per unit, is related to E_1, the wattmeter voltage, by a calibration constant k_w defined by

$$E_1=k_w\Delta P \qquad (4)$$

It should be pointed out that the input to the wattmeter is instantaneous voltage and current and ΔP is not available before the wattmeter. However, for the sake of convenience, Fig. 2 is used to derive the calibration constants.

The output of the first integrator is

$$E_2=\frac{1}{R_1C_1}\int E_1dt_1 \qquad (5)$$

and the output of the second integrator is

$$E_3=\frac{1}{R_2C_2}\int E_2dt \qquad (6)$$

To convert E_3 to the torque angle δ, it is necessary to know the conversion factor or the calibration constant k_m of the dynamic phase shifter. k_m is phase shift per volt defined by

$$\delta/E_3=k_m \qquad (7)$$

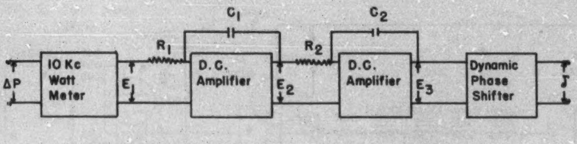


Fig. 2. Conversion of accelerating power to displacement angle

The equation for the instantaneous angular velocity of the synchronous machine is

$$\omega = \frac{180f}{HP}\int \Delta P dt \qquad (8)$$

where P is the rated output of the machine in per unit.

From equation 8, the angle δ is found by integration

$$\delta = \int \omega dt \qquad (9)$$

Comparing equations 5, 6, and 7 with equations 8 and 9, the following two relations are established

$$R_1C_1 = \frac{HPk_w\sigma}{180f} \qquad (10)$$

$$R_2C_2 = k_m\sigma \qquad (11)$$

Having calibrated the two constants k_w and k_m, and choosing a time scale σ, R_1, C_1, and R_2C_2 can be calculated from the system parameters of the problem according to equations 10 and 11, remembering that two of the four quantities may be assigned arbitrarily. Usually C_1 and C_2 are assigned convenient values, since precision capacitors without leakage and retentivity, such as polystyrene capacitors, are expensive and may not be available in all values.

Design Features

The primary frequency source of the analyzer is a conventional 10-kc crystal-controlled oscillator using two quartz-crystals of 1.01 and 1 megacycles. The output from the mixer tube is in the order of a few volts and is filtered for waveform improvement. The choice of E, I, and Z base values was made from a consideration of inductor design, since inductors are not common items of commerce. To benefit from the Iowa and Illinois designs, 2 millihenrys was selected as the per-unit inductance, so that at 10 kc the base impedance is 40 π or 125.6 ohms. With a base current of 0.1 ampere, the per-unit voltage is 4 π or 12.56 volts.

INDUCTIVE REACTANCES

In the original Iowa design, air-core coils were used. These coils were shielded in mutually perpendicular pairs. In the new Iowa and Illinois designs, inductors having molybdenum-permalloy dust cores were employed. The chief disadvantage of dust cores is that by separation of the particles an effective air gap is introduced between each particle and the over-all permeability is reduced. The ferromagnetic material known as "Ferroxcube"[9] was used in the design of inductances. This material is characterized by the possession of a high electrical resistivity together with a high magnetic permeability. Unlike dust cores, which consist of ferromagnetic particles suspended in resin binder, Ferroxcube is a homogeneous material possessing a cubic crystal structure. The high permeability of Ferroxcube and the carefully ground surfaces result in negligible external stray fields and, therefore, the cores may be placed side by side without interaction. The ability to stack these cores adjacent to each other considerably reduces the line and load units. Also a lesser amount of wire is required because of the permeability of the material.

The coils were designed in 2-millihenry and 200-, 20-, and 10-microhenry units. Their values remain constant over the current ranges in normal operation. The Q of the coils is well over 100 in all cases. This is a good feature because the resistance setting need not be shaded with such a high Q.

ZERO-IMPEDANCE AMPLIFIER

Amplifiers are used to develop the power needed in the output of the equivalent generating stations of the analyzer. Since in transient-stability studies the internal voltages of the synchronous machines are usually assumed to remain constant, the amplifiers to represent the generators should have zero output impedances. By using negative voltage feedback and positive current feedback, the effective internal impedance of the amplifier may be adjusted to zero.[10] Fig. 3 shows the circuit of the zero impedance of the amplifier. It may be noted that the number of stages must be even or a transformer must be used so that the current feedback is regenerative. The output stage consists of two power tubes connected in push-pull to a high-quality audio-frequency transformer which is stepped down in a ratio of 5 to 1. Zero effective impedance is obtained by properly adjusting the current feedback impedance. The amount of voltage feedback can be adjusted over a fair range. To vary the output voltage, it has been found more convenient to adjust the input level rather than the amount of voltage feedback. The input to the amplifier is fed from a dynamic phase shifter whose output level is readily adjustable. The amplifier output can be varied from 0.3 to 2 per unit and remain constant at the selected value up to a current of approximately 3 per unit, irrespective of the power factor.

For short-circuit studies, where the fault is close to large generating stations, it may be necessary to reduce the base voltage and base current without altering the base impedance. A potentiometer of 2 to 5 ohms in the current feedback circuit provides a voltage for feeding a vacuum-tube wattmeter.

DYNAMIC PHASE SHIFTER

The dynamic phase shifter is essentially a cascade phase-shift modulator,[11] as shown in Fig. 4. Each of the amplifier tubes *6SJ7* not only serves as a buffer amplifier to isolate the cascaded stages but also as a constant current source to its plate impedance. The plate load is a constant impedance network consisting of a coil in parallel with a series combina-

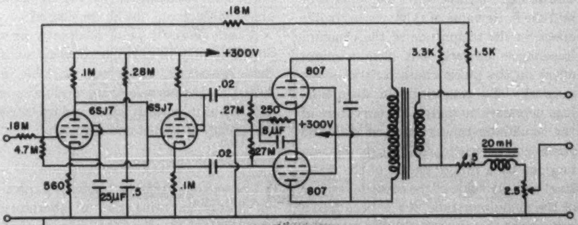


Fig. 3. Zero-impedance amplifier

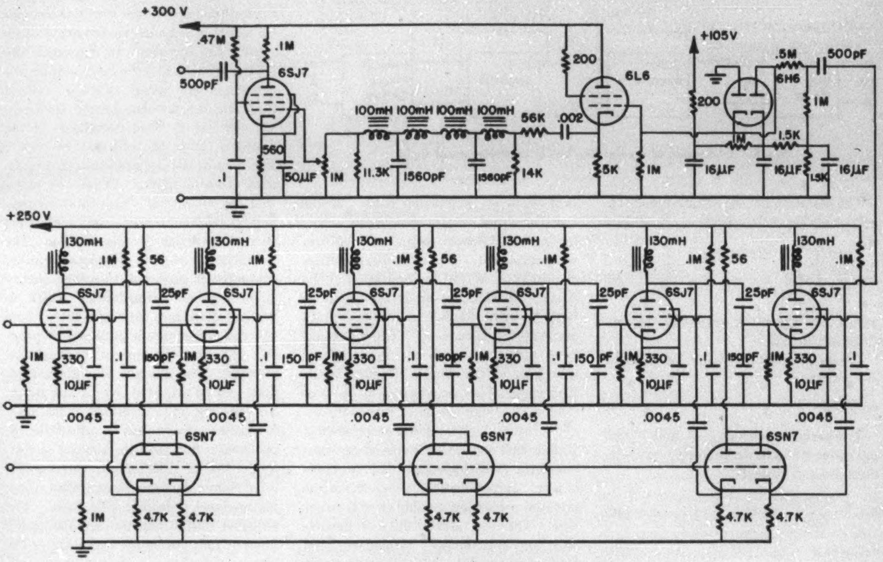


Fig. 4. Dynamic phase shifter or phase modulator

tion of a capacitor and a variable resistor. The coil has an iron-dust core with a Q of about 30 at 10 kc, while the variable resistance is presented by the output of the modulator tube *6SN7*. The function of *6SN7* is to convert a modulating signal at its control grid into a corresponding variation of resistance. When the inductive reactance of the amplifier plate circuit is adjusted equal to twice the capacitive reactance, minimum amplitude distortion will be observed at the plate. As the resistance varies in accordance with the modulating signal, the phase angle of the impedance varies, thus resulting in a phase modulation of the output signal.

Through a voltage-dividing network, the input to each amplifier tube is brought down to approximately the same level. The plate resistance of the amplifier tube and the finite value of Q have a negligible effect on the magnitude of the constant impedance network, but have a minor effect on the phase characteristic of the network. To get minimum distortion, it is necessary to match the curvature of the modulator-tube resistance with the phase characteristic. Although this can be done by a proper choice of the 10-kc level and the value of the cathode resistor of the modulator tube, it has been found convenient to adjust the bias to meet the preceding requirement. To eliminate any inherent distortion, the output of the last stage is clipped by a double diode and fed into a filter network through a cathode follower. The 10-kc fundamental is then amplified to meet the requirement of the zero-impedance amplifier. For steady-state torque adjustment, it has been found convenient to use static phase shifters preceding the dynamic ones; the static phase shifter is of the conventional design, as shown in Fig. 5. The initial steady-state phase-angle adjustment can be conveniently done by manipulating the static phase shifter until the phase angle between the signal generator and the output of the zero-impedance amplifier is the required value as shown by a cathode-ray phase-angle indicator. This is an over-all phase shift; therefore the zero-degree extra phase shift contributed by the dynamic phase shifter need not be considered.

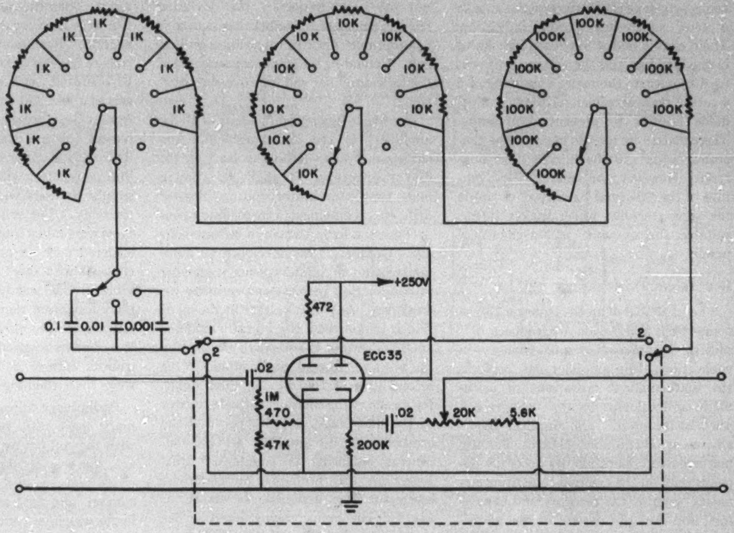
A cathode-ray type of modulator or a phasitron may be a simpler device; since these are special tubes not available in Australia, the cascade phase modulator is adopted as it requires only standard tubes and is easy to adjust and maintain.

ELECTRONIC WATTMETER

The vacuum-tube wattmeter is at present hardly anything but a laboratory instrument. It offers the only method of measuring directly the alternating wattage in very low power circuits or when high frequencies are involved. Fundamentally, it is nothing but the combination of the two rms voltmeters. Its operation depends upon a pronounced second-degree curvature in the transfer characteristic of the tubes; hence it requires carefully selected and matched tubes. The tubes *6BD7*, manufactured locally by Philips Electrical Industries, are found to be suitable for the construction of vacuum-tube wattmeters. The G_m of *6BD7*, when used as a triode, is practically constant over the range from -5.0 to 2.0 volts; thus, with a bias of -3.5 volts, the voltage applied to the tube must not exceed 1 volt rms, as the instrument is used predominantly on currents and voltages with a nearly sinusoidal wave shape.

Fig. 6 shows three *ECC35* tubes, the center one being used as a phase splitter, while the two outer ones are used as differential amplifiers. A voltage derived from the current feedback circuit in the zero-impedance amplifier is phase-splitted and then added to a voltage from the voltage divider across the zero-impedance amplifier output. The outputs from the plates of the two outer *ECC35* tubes are therefore in the form of E_1+E_2 and E_1-E_2, where E_1 is proportional to the voltage output of the zero-impedance amplifier

Fig. 5. Static phase shifter



and E_2 is proportional to the current.

The wattmeter tubes are connected in push-pull. In the plate circuit between the two tubes is a grounded wire-wound potentiometer which is used to correct for the slight difference in the characteristics of the tubes. The two equal resistances in the plate circuits are wire-wound precision resistors; their values should not be so large as to flatten the tube characteristic. The instantaneous values of the two plate voltages are

$$V_1 = A + a_1(E_1+E_2) + a_2(E_1+E_2)^2 \qquad (12)$$

$$V_2 = A + a_1(E_1-E_2) + a_2(E_1-E_2)^2$$

where A, a_1, and a_2 are the constants of the tubes, assuming the two tubes are identical.

E_1 may be written as $k_1E\cos\omega t$ and E_2 as $k_2I\cos(\omega t-\theta)$, where k_1 and k_2 are constants and θ is the angle of current lag. To get a voltage proportional to power, it is necessary to find the difference between V_1 and V_2. The result would be a d-c term proportional to $EI\cos\theta$ and terms corresponding to 10 kc and its second harmonic. The ordinary differential amplifier would not be satisfactory because slowly varying direct voltage is required and high-frequency voltages might saturate the tube. Two d-c feedback amplifiers are employed: one to reverse V_1 and the other to add V_2 and $-V_1$. The high-frequency voltages need not be filtered out since they are smoothed out by the following double integrators. The direct-current voltage corresponding to 1 per-unit power is 0.5 volt, but this can be amplified if necessary by means of the amplifier used to add V_2 and $-V_1$.

To check whether the wattmeter would give any output for zero current or zero power factor, the voltage applied to the center *ECC35* tube was reduced to zero and two equal voltages applied to the two outer tubes were varied over an appropriate range. The potentiometer in the plate circuit of the wattmeter was adjusted until the instrument gave practically zero reading over this range of voltage variation.

For transient-stability studies, the steady-state power should be balanced out before integration. This is done by feeding a voltage to the integrator equal to the steady-state power in a polarity opposite to the wattmeter indication.

Ryder and McVay[12] described a thermocouple audio-frequency wattmeter

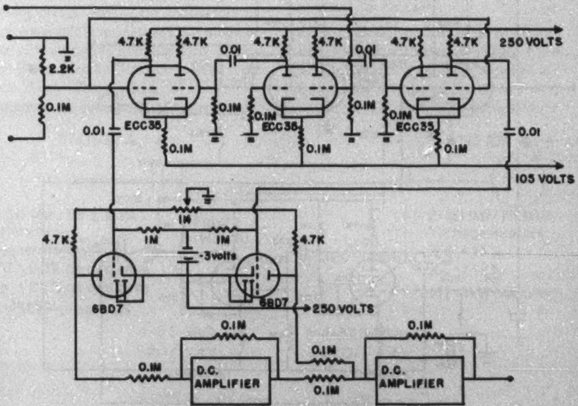


Fig. 6. Electronic wattmeter

using two identical thermocouples to give a true square-law power indication. Their circuit had a low-resistance input in the load line and the voltage drop was used to operate the meter consisting of a 4-tube 2-stage amplifier, with two double-diode limiters to prevent overloading. These meters are used to measure the generator outputs in the Illinois 10-kc analyzer. However, the inherent disadvantage of the "thermal converter" circuit is the slow response which makes it unsuitable for dynamic measurement of power.

Electronic Integrators

The basic tool of an electronic analogue computer is a high-gain d-c feedback amplifier. By mere external changes in connections, this amplifier can perform the mathematical operations of arithmetic and calculus on the voltages applied to its input. The amplifier generally has an odd number of stages. For any mathematical operation, its accuracy depends only on the external computing elements, such as the resistors and capacitors, and does not depend on the amplifier components as long as the amplifier gain remains large for all frequencies of interest. Consequently, the frequency characteristic may be used as a measure of the amplifier accuracy in producing a given operation. For instance, the frequency response of an integrator drops 6 decibels per octave. If below a certain, not too high, frequency the deviation from linearity of 6 decibels per octave is considerable, error in integration will result. Extending the frequency response of the base amplifier will improve the situation.

One of the inherent difficulties of all d-c amplifiers is the drift problem. Any change in grid-to-cathode voltage in the first stage because of a slight variation in plate supply or heater voltage, resistor drift, etc., is amplified in succeeding stages producing a large change in output voltage. Goldberg[13] has developed an automatic balancing system using a chopper in conjunction with a compensating a-c amplifier. Any drift voltage at the input grid is chopped by the vibrator and the amplifier output is half-wave rectified by the vibrator, filtered, and coupled to the second grid of the d-c amplifier input. By means of the common cathode resistor of the first stage, the chopper and the compensating a-c amplifier provide continuous balancing to counteract drift. Based on this principle, an electronic integrator was designed, as shown in Fig. 7. The compensating amplifier has a d-c gain of 1,000 and is drift-free, so that the drift voltage is reduced by a factor of about 1,000.

For stability it is important that the vibrator have make-before-break contacts.[14] Otherwise, the balancing amplifier will oscillate because of the in-phase a-c coupling through the vibrator capacitance. Siemens high-speed relays type *57a* were used for d-c amplifier drift connection. Each relay carries two separate coils of 5,000 ohms each. Only one need be operated, with about 20 to 30 volts rms of about 100-cycle-per-second supply from a Beat Frequency Oscillator. The second coil can stay unused, or may be used with a minor correction of the vibrator operation with a d-c supply. These supply voltages were fed to the base of the relay. The two leads to the two contacts were taken from the base and direct-soldered onto the brass bars carrying the contacts with short screened leads. The vibrator lead was taken to the core of the relay and from there to the base. The relays were also modified at the contacts; the vibrator carried two short platinum iridium springs and the outer contacts were both rhodium-plated.

Preliminary adjustment for the 1-to-1 mark/space ratio between the contacts, with the coil fed by its appropriate voltage, was coarsely done on the bench with signal voltages of a few volts fed to the contact, and final adjustment was made in the amplifier connections by watching 1-to-1 mark/space ratio at the a-c amplifier output and the d-c amplifier output for optimum adjustment and drift performance.

It should be pointed out that the grid current in d-c amplifiers cannot be maintained at zero as long as the plate current flows. However, it can be made practically constant by operating the tube at low plate current because the grid current is a function of the plate current. If the feedback network is a pure resistor, the grid current will develop voltages across the input resistor and the feedback resistor in accordance with the ratio of their resistance values. This is because the amplifier will drive the output voltage to buck the voltage change at the input grid, assuming that the input resistor is connected to a low-impedance source. For integrators, the grid current will charge up the feedback capacitor; it is therefore necessary to select a tube with very low grid current for the input stage. An *ECC35* or its equivalent *6SL7* tube has a grid current of the magnitude of 20×10^{-10} ampere approximately. To compensate for the grid-current flow, a small voltage from a battery can be fed to the first stage of the input tube through a very large resistance and the magnitude of this voltage is adjusted until the output voltage is zero. When two integrators are in cascade, the foregoing procedure should be applied to each individually to get a zero output for zero input during the time of the experimental run.

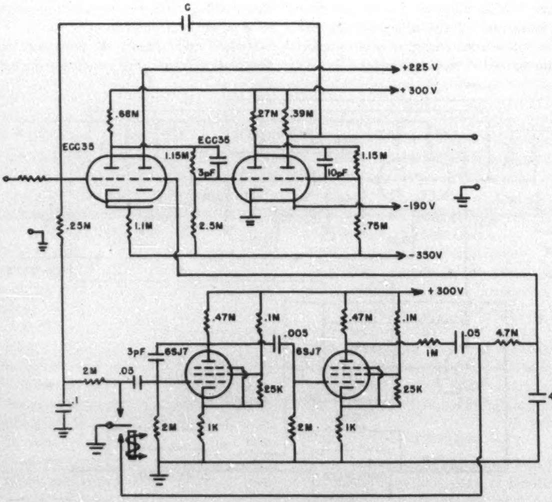
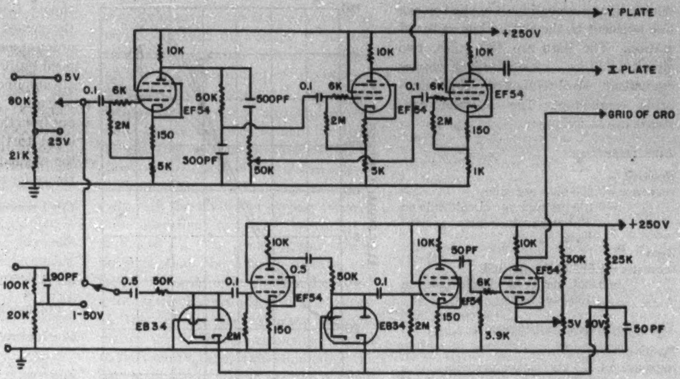


Fig. 7. Electronic integrator

Fig. 8. Phase-angle indicator


Y PLATE
+250V
X PLATE
GRID OF CRO
+250V
5V
80K
25V
21K
0.1
6K
10K
50K
500PF
2M
EF54
150
5K
300PF
50K
0.1
6K
10K
2M
5K
EF54
0.1
6K
10K
EF54
2M
150
1K
100K
90PF
20K
1-50V
0.5
50K
0.1
10K
0.5
EF54
E834
2M
150
50K
0.1
E834
2M
10K
50PF
150
6K
EF54
10K
EF54
30K
25K
3.9K
5V
50V
50PF

High-quality capacitors, preferably with polystyrene dielectric, should be used in the feedback circuit to achieve low loss and low retentivity. A switch is provided to short-circuit the capacitor of each integrator in order to return the output to voltage to zero.

Phase-Angle Indicator

The conventional ellipse method of measuring phase angle between two electrical quantities at higher audio-frequencies is not suitable in connection with this project. Because of finite spot size and harmonics, the errors become so large when the phase angle approaches 90 degrees as to make the ellipse method useless.[15] The extinguished-spot method is found to be satisfactory.[16] The principle of this method is quite simple: one 10-kc voltage is fed through an amplifier with variable gain to a phase-splitter circuit giving two equal voltages in quadrature. These two voltages are then amplified and fed to the X and Y deflecting plates of a cathode-ray oscillograph so as to produce a circular trace on the screen. Another 10-kc voltage is clipped and amfied to become a square wave. It is then differentiated and the positive pip is removed by a limiter circuit. The negative pip is fed to the grid of the cathode-ray oscillograph so as to extinguish the spot over a small portion of the circle. The position of the leading edge of the break in the circle determines the phase of the second 10-kc voltage relative to the first one. To determine the position on the circle corresponding to zero phase angle, the squaring circuit is provided with a switch so that it can also be connected to the first 10-kc voltage. A perspex scale with a number of concentric circles marked in degrees is fitted onto the screen of the cathode-ray oscillograph and is supported by small rollers so that it can be rotated by a small knob.

A complete circuit of the apparatus is shown in Fig. 8. In both channels the voltages are fed through attenuators, and it is important that there is negligible phase shift in the attenuators. This limits the input impedance to about 50,000 ohms. A cathode follower may be built outside the unit for higher input impedance, but it would limit the input voltage range. The phase shift produced in the tubes is not important as long as it remains constant, because the tubes are used to obtain both the zero setting and the reading. The size of the circle on the screen can be adjusted by varying the gain of the amplifier in the upper channel. To allow for the difference in the sensitivity of the two pairs of deflecting plates, one of the resistances in the phase-splitting network is made a potentiometer, thus enabling the X deflection to be varied until a circle is formed on the screen. The direction of spot rotation gives an indication of lagging and leading phase. This is determined and marked on the perspex scale.

The accuracy of this apparatus depends on the distortion of the wave form to be measured, since a small amount of distortion shows up on the shape of the figure obtained on the screen. The maximum possible error for a 5-per-cent third harmonic is 3 degrees. Using resistance-capacitance circuits of known values to calculate the phase angle, and comparing these with the reading taken from the apparatus, the errors are shown within 2 degrees. Small errors in the phase indicator can be corrected for, if necessary, in the experimental results. However, no accumulated error would result from this in the dynamic phase shifter and thus it is considered adequate for the determination of swing curves.

Illustrative Example

As a typical example of a transient-stability study with the use of this computer, a 2-machine problem supplied by the Electricity Trust of South Australia is given in Fig. 9, which represents a

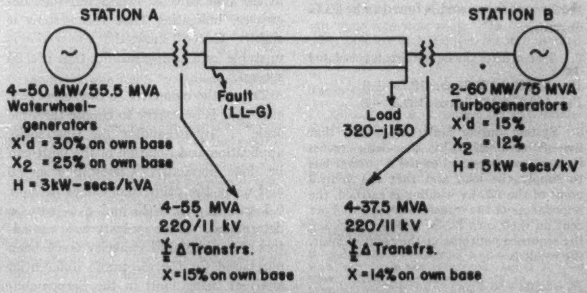

STATION A
STATION B
4-50 MW/55.5 MVA
Waterwheel-generators
X'd = 30% on own base
X2 = 25% on own base
H = 3kW-secs/kVA
Fault (LL-G)
Load 320-j150
2-60 MW/75 MVA
Turbogenerators
X'd = 15%
X2 = 12%
H = 5kW secs/kV
4-55 MVA
220/11 kV
Δ Transfrs.
X=15% on own base
4-37.5 MVA
220/11 kV
Δ Transfrs.
X=14% on own base

Fig. 9. A 2-machine problem

double-line-to-ground fault applied on one line adjacent to the 220-kv bus at the A station. The lines are 190 miles, two circuits, 220 kv, 0.5 square-inch copper equivalent steel-reinforced aluminum-cable conductors. The 50-cycle constants are as follows:

Line Impedances

Positive
sequence = 0.655 ohm per mile
= 0.135 per cent per circuit mile on 100-megavolt-ampere (mva) base
Zero
sequence = 1.587 ohms per mile
= 0.328 per cent circuit mile on 100-mva base

Shunt Capacitive Susceptance

Positive
sequence = 4.475 micromhos per mile
= 0.217 per cent per circuit mile on 100-mva base
Zero
sequence = 3.166 micromhos per mile
= 0.153 per cent per circuit mile on 100-mva base

The equivalent $-\pi$ circuit for the transmission line (positive sequence only) is:

Series impedance = 12.64 per cent on 100-mva base
Shunt susceptance of each branch = 41.5 per cent on 100-mva base

The initial operating conditions are assumed as follows:

220-kv bus voltage at $A = 107$ per cent
220-kv bus voltage at $B = 100$ per cent

Using these voltages as references, the voltages behind transient reactances of machines under normal operating conditions are $E_A = 119.9\underline{/18.44^\circ}$ and $E_B = 117.2\underline{/11.398^\circ}$. The power from A and B is given as follows:

1. System prior to fault

$$P_A = 48.31 + 235.9 \sin(19.30^\circ + \delta)$$
$$P_B = 125.83 - 235.9 \cos(109.3^\circ - \delta)$$

where δ is the displacement angle between the two machines and is found to be 20.71 degrees,

2. System after the faulty circuit is isolated

$$P_A = 31.269 + 172.34 \sin(19.70^\circ + \delta)$$
$$P_B = 143.05 + 172.34 \cos(109.3^\circ - \delta)$$

3. System during fault: Assuming that five 80-mva 22/66-kv star-delta transformers are connected to the B station bus to supply the load and that the neutral point of the 220-kv winding is earthed, the impedances of the transformers are 15.5 per cent on their own bases. After combining the sequence networks at the point of fault, the result is

$$P_A = 5.306 + 56.4 \sin(15.74^\circ + \delta)$$
$$P_B = 70.88 - 56.4 \cos(108.74^\circ - \delta)$$

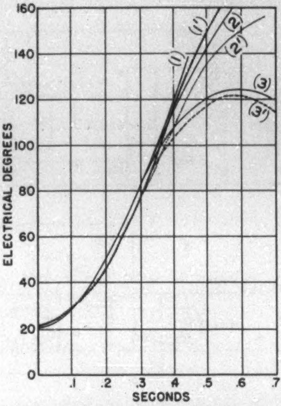


Fig. 10. Experimental and calculated results. Solid lines are calculated curves and dotted lines are experimental curves

Curves 1: Swing curves with fault applied indefinitely
Curves 2: Swing curves with the fault circuit isolated 0.3 second after the occurrence of fault, indicating that the system is stable
Curves 3: Swing curves with tripping of the faulty circuit at 0.35 second after the application of the fault, indicating that the system is unstable

The values for obtaining the swing curves are then computed and plotted as shown by the solid curves in Fig. 10. From these curves it is seen that the critical angle is 83 degrees and the time by which the faulty circuit must be cleared to maintain stability is approximately 0.31 second after the application of fault. Curve 1 shows the swing curve with fault applied indefinitely. Curve 2 is the case with faulty circuit isolated 0.3 second after the occurrence of fault. The angle reaches a maximum of 124.3 at the time of 0.58 second and then decreases, indicating that the system is stable. Curve 3 shows that the system is unstable for a fault-clearance time of 0.35 second.

To use the computer to evaluate these curves, it is necessary to choose the time scale σ. By expanding the time scale, application and isolation of the fault can be done manually. Choosing σ as 50 and using the calibrated values of $k_w = 0.5$ and $k_m = 12$, R_1C_1 and R_2C_2 can be determined. Since polystyrene capacitors of 1-microfarad capacity have been purchased for the electronic differential analyzer being built in the Servomebhanisms Laboratory, it was decided to choose $C_1 = C_2 = 1$ microfarad. To keep the values of R_1 and R_2 about the same order of magnitude, equation 10 is multiplied by 10^{-2}, thus maintaining the over-all amplification of the two integrators constant. However, by doing so, the output from the first integrator and hence the input to the second one corresponds to the angular velocity multiplied by 10^{-2}. Hence for machine A

$C_1 = 1$ microfarad; $C_2 = 1$ microfarad

$$R_1 = \frac{3 \times 2.22 \times 0.5 \times 50}{180 \times 50} \times 10^2 = 1.85 \text{ megohms}$$

$$R_2 = 12 \times 50 \times 10^{-2} = 6 \text{ megohms}$$

For machine B

$C_1 = 1$ microfarad; $C_2 = 1$ microfarad

$$R_1 = \frac{5 \times 1.5 \times 0.5 \times 50}{180 \times 50} \times 10^2 = 2.08 \text{ megohms}$$

$$R_2 = 6 \text{ megohms}$$

The experimental results, after having been converted to the actual time scale, are plotted as dotted curves in Fig. 10. The experimental curves follow faithfully the calculated curves up to approximately 0.35 second and then begin to depart. This is caused by slight drifts in the wattmeter tubes. On the whole, the two sets of curves compare favorably and conclusions can be drawn as to whether the system is stable or not from the computer solutions.

The calculated curves were made by using time intervals of 0.05 second, so it may be expected that there should be cumulative errors of the stepwise technique for large values of t. The computing solutions, on the other hand, are continuous and were recorded manually and therefore are subject to human errors. A Sanborn recorder or other types of automatic recorder having appropriate frequency response can be used to eliminate the human errors. One inherent disadvantage of the electronic wattmeter is the slight drift which cannot be completely eliminated. It is doubtful if it can even be minimized simply by a proper choice of time scale because, if the time scale is compressed, it is necessary to increase the over-all amplification of the feedback loop proportionately.

Built as a pilot model to test its performance, plans have been made to develop more suitable wattmeters for 10-kc frequency in order to analyze such problems as a study of the stability during subsequent swings of a system or the pulling into step of the elements of a system. One is to modify the wattmeter now being used in the master measuring desk of the Illinois Analyzer.[17] This is a Weston model *432* with a fixed coil rating of 750 milliamperes and a moving coil current rating of 13 milliamperes for full deflections. A suitable voltage amplifier and a current amplifier are used in connection with this wattmeter. To eliminate the mutual reactance between the two coils and the distributed capacitance of the coils, a conventional bridge circuit with one of the meter coils in one of the bridge arms is connected to the output of one of the amplifiers. A feedback circuit is arranged such that, if a mutual voltage is induced in the coil, a portion of this voltage is fed back into the amplifier in such a way as to cancel the effect of this error voltage on the meter coil current. The correction voltage is independent of the desired driving voltage, i.e., the function of the bridge circuit is to separate the desired driving voltage from the error voltage. The movement of this meter may then be used to vary a potentiometer and hence give a voltage proportional to the wattmeter reading. In this case, drift is entirely eliminated.

Another method being explored originates from a project on the development of a high-speed multiplier utilizing the fact that the discharge or charge of a capacitor through a resistance follows an exponential law. Assuming that two identical resistance-capacitance networks initially charged to V_1 and V_2 are allowed to discharge through resistors at $t=0$, then at $t=t_1$ V_1 falls to V_3 while V_2 falls to a value $(V_2V_3)/V_1$. Taking V_1 as reference and sampling the voltage of the second resistance-capacitance network at the instant when the voltage of the first resistance-capacitance network reaches the level V_3, the product of two quantities is obtained. A switching circuit recharges the resistance-capacitance networks to new levels of V_2 and V_3; then the charging circuits are switched off to allow normal discharge until a new level of V_3 is reached. This is done automatically with the multiplier cycling through these operations to follow the variation of V_2 and V_3. The sampling rate should be such that the answer storage can follow the variation i.e., within the period t_1 the inputs should undergo little changes. In this scheme the accuracy depends mainly on the circuit elements external to the vacuum tubes, thus eliminating the disadvantages of previous methods which depend entirely on the tube characteristics. After this scheme is perfected, it can be made to work as a wattmeter free from drift.

References

1. An Electric Analogue Method for the Direct Determination of Power System Stability Swing Curves, **W. B. Boast, J. D. Rector.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1833–35.

2. Model Machines and Model Systems: Study of Transient-Stability Problems by the Use of Electromechanical Models Resembling the Existing Machines and Systems, **R. Roberts.** *Paper No. 338*, CIGRE, Paris, France, 1950.

3. *Discussion* by **Alexander Kusko** of reference 1, pp. 1835–36.

4. A High Frequency Simulator for the Analysis of Power Systems, **S. Kaneff.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 100, pt. II, Aug. 1953, pp. 405–16.

5. Analysis of Hunting Phenomena of Power Systems by Means of Electrical Analogues, **D. W. C. Shen, J. S. Packer.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 101, pt. II, 1954, pp. 21–34.

6. Analyzer Interconnections for Direct Determination of Power System Swing Curves, **G. A. Beckey, F. W. Schott.** *AIEE Transactions*, vol. 73, April 1954, pp. 238–42.

7. Design Improvements and Operating Experience with 10-Kc Network Analyzers, **J. D. Ryder, W. B. Boast.** *AIEE Transactions*, vol. 71, pt. I, 1952 (Jan. 1953 section), pp. 437–42.

8. Power System Stability (book), **S. B. Crary.** John Wiley and Sons, Inc., New York, N. Y., vol. II, 1947.

9. Ferroxcube Materials, **C. M. Van Der Burgt, M. Gevers, H. P. J. Wigin.** *Philips Technical Review*, Eindhoven, Holland, vol. 14, March 1953, pp. 245–56.

10. Control of the Effective Internal Impedance of Amplifiers by Means of Feedback, **H. F. Mayer.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 27, 1939, pp. 213–17.

11. Cascade Phase Shift Modulator, **M. Marks.** *Electronics*, New York, N. Y., vol. 19, Dec. 1946, pp. 104–09.

12. A Thermocouple A. F. Wattmeter, **J. D. Ryder, M. S. McVay.** *Radio-Electronic Engineering*, Feb. 1953, pp. 6–7.

13. Stabilization of Wide-Band Direct-Current Amplifiers for Zero BDN Gain, **E. Goldberg.** *RCA Review*, Princeton, N. J., June 1950, p. 296.

14. Driftless D. C. Amplifiers, **F. R. Bradley, R. McCoy.** *Electronics*, New York, N. Y., April 1952, pp. 144–48.

15. Phase Angle Measurements, **F. A. Benson, H. O. Carter.** *Electronic Engineering*, London, England, vol. 22, 1950, p. 238.

16. A Versatile Phase-Angle Meter, **G. N. Patchett.** *Electronic Engineering*, London, England, vol. 24, 1952, pp. 224–29.

17. Report on 10 Kc Wattmeter, **Network Analyzer Laboratory.** University of Illinois, Urbana, Ill., 1953.

◆

Discussion

G. A. Bekey (University of California, Los Angeles, Calif.): This paper presents an interesting extension of several recent proposals for automatic determination of system swing curves. The different approaches to this problem may be listed as follows:

1. Construction of special-purpose analogues: A number of such analogues have been constructed; see references 1 through 5 of the paper. The main disadvantage of this type of device is that it is specialized in function and may require network analyzer data on the power system under consideration.

2. A step-by-step swing curve computer: A device of this kind has now been added to the Westinghouse network analyzer at East Pittsburgh[1] and to the General Electric network analyzer in Schenectady; this swing curve calculator is shown on the May 1954 cover of *Electrical Engineering*. This device is simply a small computer which carries out the slide-rule calculation for each angle increment when the power output of each generator is fed into it and displays this angle as a voltage.

3. Automatic determination of swing curves using existing network analyzers: In this third category fall devices such as that described by Mr. Shen and Mr. Lisser which make use of a network analyzer on which the steady-state behavor of the system may also be studied. A computer and servomechanism are used with each generator. The authors present a discussion dealing with the 10-kc analyzer. Another fully automatic scheme suggested by Dr. E. L. Harder of Westinghouse Electric Corporation was built by Giovànni Malaman at Yale University.[2,3] A similar scheme made use of a mechanical differential analyzer in conjunction with a 480-cycle network analyzer (see ref. 6 of the paper).

4. Construction of new synchronous machine analogues: It is also possible to simplify the study of transient stability by constructing a new analogue for the synchronous machine to be used with the network analyzer. Such analogues which give the quadrature- and direct-axis voltages and currents explicitly have been constructed by J. E. Van Ness.[4]

The greatest stumbling block to ready acceptance of automatic methods such as the one described by the authors seems to be the lack of accurate, trouble-free components. Devices such as electronic wattmeters and phase shifters still present problems in accuracy and reliability. The present paper is another step in the direction of eliminating these problems.

It is probably still too early to determine which of the four approaches I have listed will come into wide usage. In any case, it is my opinion that the step-by-step method will soon be eliminated.

REFERENCES

1. *Discussion* by **R. B. Squires** and **E. L. Harder** of reference 6 of the paper, p. 241.

2. Undispostive per lo Stúdio dei Sistemi Electtrici in Regime Perturbato, **Giovanni Malaman.** *Thesis*, Yale University, New Haven, Conn., 1952.

3. Possible Improvements to A-C Network Analyzers, **Giovanni Malaman.** *Societa Edison*, Milan, Italy.

4. Synchronous Machine Analogues for use with the Network Analyzer, **James E. Van Ness.** *AIEE Transactions*, vol. 73, pt. III, Oct. 1954, pp. 1054–60.

Stabilized Power Supplies for Instrument Applications

W. G. AMEY
MEMBER AIEE

F. H. KRANTZ
ASSOCIATE MEMBER AIEE

W. R. CLARK
MEMBER AIEE

A. J. WILLIAMS, JR.
MEMBER AIEE

IN SOME instrument applications, batteries are used as sources of d-c power because they are portable, and free of internal noise (except for slow changes with temperature and time), and because they can be completely shielded from external noise. These advantages are frequently offset, however, by problems of size, weight, and replacement, and as a result increasing use is being made of a-c powered electronic d-c power supplies for a-c amplifiers, d-c amplifiers, and d-c potentiometers for voltage measurement. Although such electronic d-c supplies require a-c input power, this requirement is generally not objectionable because a-c power outlets are conveniently available. Means for suppressing the effects of line voltage fluctuation are often required for a-c amplifiers, generally required for d-c amplifiers, and always required for d-c potentiometers for voltage measurement.[1] The use of a series (rheostat) tube with an appropriate d-c control amplifier and voltage reference has been generally accepted as a preferred way to accomplish this,[2] an example of which is shown in Fig. 1. The heavy lines indicate the path of the power to the load R_L. The series rheostat tube, under the control of the components indicated by light lines, stabilizes the d-c load voltage V_L.

This paper is concerned with: 1. the voltage reference or voltage standard; 2. zero drift in the d-c control amplifier; 3. means of checking and correcting for this drift; 4. adjustment of load voltage; 5. radio-frequency (r-f) noise; 6. making voltage and voltage stability adequate for photomultiplier tubes. When necessary, elaborate means of coping with these problems can be used to provide precisely known output voltages with exceptional stability. The relatively simple means described in this paper are adequate for many instrument applications. (It is not intended to imply that every item described is commercially available.)

Factors Affecting Stability

Fig. 1 indicates schematically the basic type of stabilized system under consideration. The power transformer, rectifier, and filter convert the unregulated a-c input to unregulated direct current at a suitable high voltage level. The unregulated direct current is converted to regulated voltage V_L across the output load resistance R_L by means of the series power tube used as a rheostat. The resistance of the rheostat tube is adjusted by suitably varying its grid potential in accordance with deviations in the voltage V_L across R_L.

To reduce such output deviations, they are fed through a suitable amplifier to the grid of the rheostat tube. The electronic control system is fast enough to stabilize V_L against rapid changes in either load current or a-c supply voltage.

Deviations of V_L from the desired normal value are detected by comparing a portion of V_L with a stable voltage reference, the net difference being fed to the d-c control amplifier. As has been indicated previously[2-4] such an arrangement effectively constitutes a feedback amplifier. When the loop gain is sufficiently high, the output voltage is nearly as stable as the voltage reference, despite variations in the load current, the unregulated d-c voltage, and the control amplifier gain. Techniques for providing adequate loop gain (without oscillation) have been given.[2-4] In the final analysis, however, the output voltage stability is determined by:

1. Deviations of the feedback factor β.
2. Deviations of the voltage reference.
3. Control-amplifier-zero deviations, which are equivalent in effect to deviations of the voltage reference.

For many instrument applications, these three principal deviations must be so controlled or reduced that satisfactory over-all stability is maintained for periods of 8 hours or more. Deviations in β are readily reduced by suitable arrangements of stable resistors. Even though the resistance values may change markedly with temperature, the resistors can usually be so matched as to minimize the effect on β. Unfortunately, deviations in the voltage reference and control amplifier zero are not so readily avoided.

Voltage Reference and D-C Control Amplifier

The electrochemical standard cell can be used for the voltage reference but it is not adapted to a wide temperature range. Its relatively small voltage appears to dictate the use of a chopper amplifier[5] or a chopper stabilized d-c amplifier.[6] Highly evacuated tungsten filament lamps

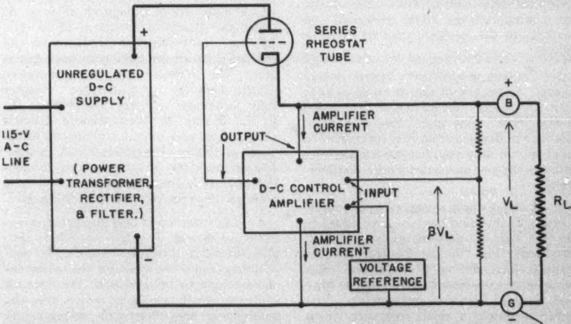


Fig. 1. An example of an electronically stabilized system

Paper 54-295, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Summer and Pacific General Meeting, Los Angeles, Calif., June 21-25, 1954. Manuscript submitted March 23, 1954; made available for printing April 28, 1954.

W. G. Amey, F. H. Krantz, W. R. Clark, and A. J. Williams, Jr., are with the Leeds & Northrup Company, Philadelphia, Pa.

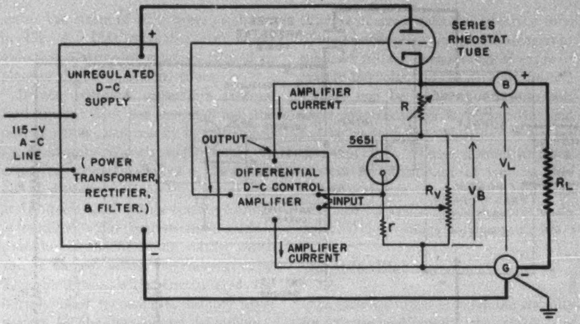


Fig. 2. System using voltage-sensitive bridge. There is only one voltage V_B for which the bridge is balanced. For the balanced condition there can be only one value of current into the bridge

in a relatively high-voltage bridge were found stable in tests but the bridge attenuated the output-voltage change as seen by the control amplifier, the requirements on which were thus made more stringent. A bridge using a single glow-discharge voltage-regulator tube, as shown in Fig. 2, introduces less attenuation, but such tubes available prior to 1946 were somewhat subject to instabilities of a spontaneous nature.[7] Fortunately the *S-901-C* glow tube was supplied to us in 1946 by Radio Corporation of America. This development tube,[8] and the commercial tube *5651* which followed, are free of the instabilities of a spontaneous nature found in the earlier tubes such as the *VR105/30*.

Experience with *5651* tubes to date has indicated that, with very few exceptions, they also exhibit remarkably good characteristics with respect to both long-time stability of operating voltage and repeatability of operating voltage with successive tube firings. The temperature coefficient for such tubes is between −0.003 and −0.005 per cent per degree centigrade over very wide ranges of ambient temperature. After exposure to severe thermal shock, they exhibit no significant changes in their characteristics. They have an incremental resistance of between 1,000 and 1,500 ohms. This is considerably higher than that of several other types, but the small operating current of 1.5 to 3.5 milliamperes (ma) makes it tolerable. The stability of the *5651* is so good that a "differential" d-c control amplifier with its reduced zero shift[9,10] is justified.

A Power Supply Based on 5651 Tube and Differential D-C Control Amplifier

Prior to World War II, a recording microphotometer was developed to measure and record the relative densities of lines on spectrographic plates.[11] This instrument required two separate d-c supplies having outputs stable to better than 0.1 per cent for periods of 8 hours or more. One supply was needed for a low-level conductively coupled d-c amplifier; the other was needed for the special lamp by means of which the photographic plate was scanned. None of the then commercially available regulated supplies were found to be suitable for this application, and d-c power supplies had to be designed. Early designs were based on the *VR105/30* in a circuit described later. Because some of these tubes changed voltage in an abrupt and spontaneous way, as previously mentioned,[7] elaborate long-time tests were required. These difficulties were known to the tube experts, who informed us that the voltage jumps could not readily be cured in the *VR105/30* tube and recommended redesigning the d-c supplies around the new tube which should be free of the voltage jumps.[8] This was done based on the system shown in Fig. 2. The new unit has a nominal output of 105 volts, and the output voltage can be set precisely at 105 volts by adjustment of the tap on R_V.

Deviations of V_L from a normal value are detected and fed to the control amplifier by means of the voltage-sensitive bridge shown in Fig. 2. This bridge consists of the two portions of R_V, the fixed resistor r, and the *5651* gaseous discharge tube. Since the voltage drop across the *5651* tube does not vary with current as do the resistive voltage drops in the other arms, the bridge is balanced for only one value of the voltage V_B applied to its input. As shown in Fig. 2 the d-c control amplifier is the differential type, both of its signal input terminals being grids. Since these grids do not draw appreciable current, zero drift of the control amplifier current-wise is negligible. Voltage-wise, the zero-drift of the differential control amplifier is small because of symmetry.[9,10] Although the differential amplifier does not completely eliminate heater-voltage effects, the compensation appears to be adequate without special tube selection and without other means of compensation. When the a-c line voltage is changed abruptly from 100 to 125 volts, the control amplifier quickly acts to limit the output voltage change to 0.025 per cent. Within a few seconds, however, the unbalance in the heaters (which have been subjected to a proportional variation in a-c supply) causes an extra change of about 0.025 per cent in output voltages. Depending on the direction of heater unbalance for the particular tube combination, such change may add to or subtract from the initial 0.025-per-cent variation. In other words, the heater effect may bring about almost perfect compensation or it may cause the total output variation to approach 0.05 per cent. Regulating the power to the heaters could do no more than hold this total variation to about 0.025 per cent (as set by the gain of the d-c control amplifier). Further advantages of the differential control amplifier are described later.

Utilization of Voltage-Sensitive Bridge for Adjustment of Load Voltage

With R (Fig. 2) set at zero, the load voltages V_L is the same as the bridge balance voltage V_B. This voltage can be precisely adjusted to any value within a range of a few per cent above or below the nominal 105-volt value by adjustment of the tap point on R_V. This type of adjustment, limited in range by the allowable limits of current through the *5651* tube, is intended primarily to restore V_L to the desired nominal value when a *5651* tube is replaced by one having a slightly different operating voltage. Whenever V_L need not be set to a specific value, the tap point on R_V need not be made adjustable, and two fixed resistors can be used. The load voltage will then be stable, but its value will depend on the particular *5651* tube in use. The fixed ratio and the resistor r in series with the tube must be assigned such values that none of the *5651* tubes used will be oper-

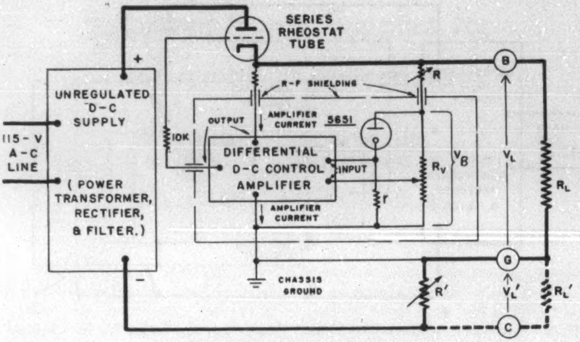

SERIES RHEOSTAT TUBE
+
UNREGULATED D-C SUPPLY
115-V A-C LINE
(POWER TRANSFORMER, RECTIFIER, & FILTER.)
−
10K
R-F SHIELDING
OUTPUT
AMPLIFIER CURRENT
5651
DIFFERENTIAL D-C CONTROL AMPLIFIER
INPUT
AMPLIFIER CURRENT
r
R
R_V
V_B
V_L
R_L
B
G
C
CHASSIS GROUND
R′
V_L'
R_L'

Fig. 3. Voltage adjustments and r-f shielding

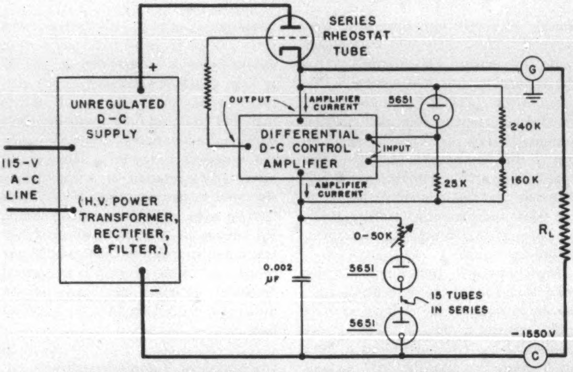

SERIES RHEOSTAT TUBE
+
UNREGULATED D-C SUPPLY
115-V A-C LINE
(H.V. POWER TRANSFORMER, RECTIFIER, & FILTER.)
−
OUTPUT
AMPLIFIER CURRENT
5651
DIFFERENTIAL D-C CONTROL AMPLIFIER
INPUT
AMPLIFIER CURRENT
G
240K
25K
160K
R_L
0-50K
0.002 µF
5651
15 TUBES IN SERIES
5651
−1550V
C

Fig. 4. Simplified circuit for photomultiplier supply. The output voltage can be reduced by reducing the variable 50K resistor and by short-circuiting one or more of the 5651 tubes

ated outside the allowable range of tube current (1.5 to 3.5 ma).

Whether the tap on R_V is adjustable or fixed as just noted, the control amplifier tends to maintain the bridge voltage at its balanced value V_B. As long as the voltage-current characteristic of the *5651* tube remains unchanged, the total balanced bridge current will be constant (if the resistance of *r* and the total resistance of R_V are constant). If *R* is now set to a resistance value other than zero, the power supply output voltage V_L exceeds V_B by the amount of the voltage drop acrosss *R*. Since the bridge current as well as the bridge voltage is stabilized by the control amplifier, V_L is stabilized (if *R* does not vary). The difference between V_L and V_B increases linearly with increase in *R*.

When V_L is increased by means of *R*, the plate supply of the d-c control amplifier increases also. Because the amplifier is of the differential type, the only significant effect of such change is a slight increase in amplifier gain. This tends to compensate automatically for the attenuation introduced by *R* in the feedback path. In fact, experiments show that the resultant stability is slightly better with V_L at 150 volts than it is with V_L at 100 volts.

The microphotometer[11] previously mentioned required a negative voltage in addition to a well-stabilized positive voltage. Fig. 3 shows how this is done by returning the negative side of the unregulated d-c supply to ground through resistor R' instead of through a direct connection. Assuming R_L' infinite, the total current through R' consists of the external load current through R_L, the bridge current, and the current drawn by the differential control amplifier. The amplifier current is essentially constant because the differential action tends to increase the current in either tube of a given stage whenever the current decreases in the other tube of that stage. Consequently, so long as R_L is held constant, the total current through R' remains essentially constant. If R' is constant, a relatively well-stabilized voltage, negative with respect to ground, is thereby provided, which increases or decreases with R'. When an external load R_L' is connected from the negative terminal *C* to ground, the negative voltage decreases in accordance with the equivalent resistance of R' and R_L' in parallel. For applications where both R_L and R_L' are essentially constant, the value of R' can be adjusted to provide the desired negative voltage.

Versatility of New Power Supply

The useful features of the voltage-sensitive bridge, in combination with those of the differential control amplifier, have made it possible to provide for output voltage adjustment over a range from 100 to 150 volts without significant loss of stability. At 105 volts' output, load current change from zero to 200 ma produces only a 0.05-per-cent change in output voltage. At a 150-volt output, the allowable maximum load current is reduced to about 150 ma. At any load currents within ranges mentioned, change in a-c line voltage from 100 to 125 volts changes the d-c output voltage only 0.05 per cent. At constant load, with ordinary a-c line voltage fluctuations, the long-term drift in output voltage after a 20-minute warm-up is not more than 0.05 per cent in 24 hours.

Effects of Radio-Frequency Noise

In many industrial installations, the a-c power lines pick up and conduct substantial amounts of r-f noise originating from commutation, switching, etc. Because the tubes are fast and nonlinear their output can be shifted by a relatively small amount of such noise. In brief, the tubes act as r-f detectors. To avoid this the critical tubes should be shielded from r-f noise. Fig. 3 shows schematically how such arrangements were made in one critical application. Since the voltage-sensitive bridge has several connections to the d-c control amplifier, they are both placed in the same shield box so that only a few mica feed-through capacitators are required to provide r-f filtering.

High-Voltage Power Supply with Stability Adequate for Photomultiplier Tubes

Because the output of a photomultiplier tube may vary as the seventh power of the d-c voltage the stability requirements are stringent for some instrument applications. The techniques previously cited can be applied in a high-voltage power supply with adequate stability, as illustrated in Fig. 4. The 2-stage differential control amplifier receives its input signal from the voltage-sensitive bridge formed by the *5651* tube and the three fixed resistors. With these fixed resistance values, the voltage across the bridge depends on the particular *5651* tube used, but the tube current is never outside the allowable range of 1.5 to 3.5 ma. In fact, in accordance with both calculation and experiment, the total of the bridge current and control amplifier current always lies in the range 2.9 to 3.1 ma. For any given *5651* tube used in the bridge, this total current is held essentially constant by the control amplifier, and has just the right value for the string of series connected *5651* tubes through which it is passed.

The string of *5651* tubes is equivalent to a single high-voltage reference tube. Just as in the use of *R′* in Fig. 3, these tubes, with constant current through them, provide a constant voltage drop which is added to the bridge voltage to obtain the total load voltage of the power supply.

The external load current can have any value consistent with the limitations imposed by the unregulated d-c supply, the rheostat tube, and the control amplifier. For the arrangement of Fig. 4, variation of the external load from zero to 3 ma produces less than 0.01 per cent change in load voltage. At constant load resistance, the load voltage varies less than 0.01 per cent for an a-c line voltage change from 100 to 125 volts. After about 20 minutes' warm-up, at constant load current and normal a-c line fluctuations, the load voltage is stable to within about 0.01 per cent for periods of 8 hours or more.

The string of *5651* tubes closely couples the over-all load voltage deviations to the voltage-sensitive bridge, thereby reducing the gain requirements for the control amplifier. Resistors can be used in place of the *5651* tubes if one can tolerate the reduction in output stability which results from the higher coupling impedence and the temperature coefficient of the resistors. To avoid oscillation, a 0.002-microfarad capacitor is necessary across the string of *5651* tubes as shown in Fig. 4. This capacitor also is required if a string of resistors is used in place of the string of *5651* tubes.

If the required maximum external load current does not exceed 3 ma, the output voltage can readily be reduced in approximately 87 volt steps simply by using fewer reference tubes in the string. No modification of the differential control amplifier and voltage-sensitive bridge is required. The limitation on such a procedure is the maximum voltage which can be allowed across the rheostat tube. This limitation when encountered has been avoided by using a suitably lower voltage for the unregulated d-c supply.

The adjustable 50,000-ohm rheostat (Fig. 4) in series with the reference tube string provides a means for adjustment of the load voltage over a range of approximately 150 volts. By this means the load voltage can be adjusted either to suit one's convenience or to restore the load to the desired level with any given set of *5651* tubes. These general arrangements for adjustment of load voltage afford a degree of simplicity and convenience not available in other previously reported high-voltage stabilized supplies.[12]

Quest for Stability

In striving to get comparable stability for smaller load voltage, the zero drift of the d-c control amplifier becomes more of

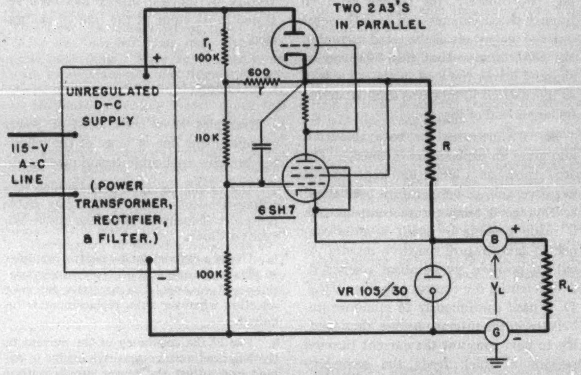


Fig. 5. Early circuit for cascaded system. The regulator tube across the load attenuates the zero-drift of the control amplifier

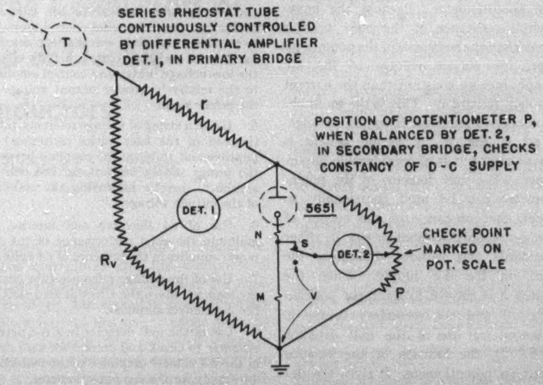


Fig. 6. An arrangement for checking the constancy of a stabilized d-c supply for a potentiometer. The resistor N with a positive temperature coefficient compensates the small negative temperature coefficient of the 5651 tube

a problem. An early attack on this problem, used for the microphotometer,[11] is shown in Fig. 5 where the glow tube is placed directly across the constant load resistance. In this position the constant-voltage characteristic of the glow tube attenuates the zero-drift in the control amplifier. Instability in the *VR105/30* glow tube led to the discard of this particular arrangement for the microphotometer but a similar attack with improved components and techniques has been used in an arrangement to provide current for d-c potentiometers for voltage measurement. For such measurements, the current through the potentiometer must be held constant with change of conditions and with lapses of time at least as long as the period between checking. Fortunately, the current required by such potentiometers can be of the same order of magnitude as the rated current of the *5651* tube so that this tube can be shunted across the load in the same way as the *VR105/30* tube was shunted across its larger load of Fig. 5.

Fig. 6 illustrates the bare essentials and gives an explanation of them. The *5651* tube has a small and reproducible negative voltage-temperature coefficient[7] so that good temperature compensation is made possible by using a series-connected temperature sensitive resistor *N*, with a positive temperature coefficient. A differential d-c control amplifier (*DET* 1), is used continuously to minimize unbalance in the primary bridge, thus tending to hold constant the current through resistor *r* which feeds the secondary bridge. The voltage across the secondary bridge is the output voltage of the stabilized d-c supply and the input voltage of the potentiometer. Because the incremental resistance of the *5651* tube is lower than the resistance of the potentiometer, the output voltage of the d-c supply is more constant than the current through resistor *r*. This is the more obvious benefit derived from the arrangement of Fig. 6. A second benefit is less obvious but is of considerable importance. The null detector of the potentiometer can be used occasionally to check the constancy of the output of the d-c supply without resort to other standards and without disturbing the load on the d-c supply. When the switch *S* is thrown to its upper position (Fig. 6) and the secondary bridge is balanced by use of the null detector (*DET* 2), the location of the balance point on potentiometer *P* gives the desired check. If the position of this balance point has not changed from its position at the time of calibration, the output voltage of the d-c supply has not changed since the time of calibration. This assumes constancy in the elements in the secondary bridge but does not assume constancy of anything else.

Because the incremental resistance of the *5651* tube is approximately known, the change in the output voltage of the d-c supply can be evaluated from the change in the position of balance for potentiometer *P*. In this way, a change of 0.1 per cent in the output voltage could readily be detected if it should occur. If the output voltage of the d-c supply does change, it can be restored to its initial value by adjusting slowly one of the resistive arms of the primary bridge. When the potentiometer *P* again balances at the calibration mark, the output voltage of the d-c supply has been restored to its value at the time of calibration.

Conclusion

Examples have been cited of power supplies using one or more of the following broadly applicable techniques:

1. Use of the voltage reference tube in a voltage-sensitive bridge with highly stable resistance elements to provide a stable balance condition.
2. Use of a differential d-c control amplifier to reduce the effect of heater voltage variations and to avoid the necessity for tube selection wherever tube replacement is required.
3. Use of the constancy of the current to the balanced voltage-sensitive bridge to extend and adjust the power supply output voltage by means of a simple rheostat in series with the bridge.
4. Use of the constancy of the current to the parallel combination of the balanced bridge and the differential control amplifier to provide a stabilized voltage drop across a fixed series resistance, and thereby couple the low-voltage bridge and control amplifier to the relatively greater output voltage of the power supply.
5. Use of a string of voltage reference tubes (instead of the fixed series resistance) to improve and stabilize the coupling between the power supply output and the control amplifier, thereby improving the stability of the output voltage.
6. Use of r-f shielding and filtering to maintain the good performance of the d-c power supplies in the presence of r-f noise.
7. Use of the voltage reference tube across the load to reduce the effect of zero shift in the d-c control amplifier.
8. Use of the null detector in a d-c potentiometer to check and correct for zero drift in the d-c control amplifier in the stabilized d-c supply for the d-c potentiometer.
9. Use of a resistor of positive temperature coefficient to compensate for the negative temperature coefficient of the *5651* tube.

Stabilized power supplies using these techniques have characteristics well suited to the stringent requirements of many instrument applications.

References

1. Voltage Regulating System, **A. J. Williams, Jr.** United States Patent No. *2,032,455*, Mar. 1936.

2. Regulating Device, **A. W. Vance.** United States Patent No. *2,075,966*, April 1937.

3. Electronic Voltage Stabilizers, **F. V. Hunt, R. W. Hickman.** *Review of Scientific Instruments*, New York, N. Y., vol. 10, Jan. 1939, pp. 6–21.

4. Voltage-Regulated Power Supplies, **Leonard Mautner.** *Electrical Engineering*, vol. 66, Sept. 1947, pp. 894–900.

5. D.C. Amplifier Stabilized for Zero and Gain, **A. J. Williams, Jr., R. E. Tarpley, W. R. Clark.** *AIEE Transactions*, vol. 67, pt. I, 1948, pp. 47–57.

6. Wide-Band D-C Amplifier Stabilized for Gain and for Zero, **A. J. Williams, Jr., W. G. Amey, Will McAdam.** *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 811–15.

7. Characteristics of Certain Voltage-Regulator Tubes, **George M. Kirkpatrick.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 35, May 1947, pp. 485–89.

8. Voltage Reference Tube, **Gerald G. Carne.** United States Patent No. *2,556,254*, May 15, 1947.

9. A High-Gain D-C Amplifier for Bio-electric Recording, **Harold Goldberg.** *AIEE Transactions*, vol. 59, Jan. 1940, pp. 60–64.

10. Cathode Phase Inversion, **O. H. Schmitt.** *Journal of Scientific Instruments*, London, England, vol. 15, 1938, pp. 100–01.

11. A New Recording Microphotometer, **R. C. Machler.** *Proceedings*, 7th Summer Conference on Spectroscopy and Its Applications, Technology Press, Massachusetts Institute of Technology, Cambridge, Mass., pp. 65–67, July 1939.

12. A Stable Source of High Voltage, **L. V. Hibbard, D. E. Caro.** *Journal of Scientific Instruments*, London, England, vol. 30, Oct. 1953, pp. 378–80.

Discussion

Byron M. Jones (Reliance Electric and Engineering Company, Cleveland, Ohio): The voltage sensitive bridge and the differential control amplifier discussed in this paper certainly offer the possibility of better voltage regulation than would be possible with a conventional voltage reference tube and d-c amplifier feedback system. The constant current characteristic of both the bridge and the differential amplifier should prove useful in a variety of circuit arrangements. The authors are certainly right that the type-*5651* voltage reference tube is a great improvement over the older voltage regulator tubes.

The following remarks concern photomultiplier power supplies. While working on an undergraduate thesis at Purdue University (concerning a fast coincidence circuit for nuclear research), the writer needed a power supply that was more stable than 0.01 per cent. Such circuits are described in Higinbotham's paper,[1] but some circuit simplifications can be made by not using a d-c feedback system. The principle load on a photomultiplier power supply is the voltage divider across the dynodes of the tube. Since this is a constant load, the power supply need not regulate for wide

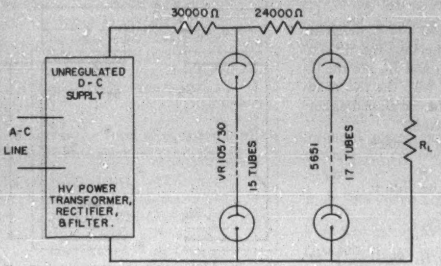


Fig. 7. Two-stage regulator circuit

load current changes. One characteristic of the d-c feedback amplifier is that it expands the operating range of the voltage reference tube If the operating range as dictated by the load characteristic is not too great, there is little reason for using the feedback.

The circuit shown in Fig. 7 was first used by the Atomic Energy Commission at Oak Ridge, Tenn. At Purdue University it was modified slightly. The principle of the operation of the circuit is as follows. The first string of voltage regulator tubes (type *VR 105/30*) bring voltage to a semistable value. The second string of voltage reference tubes (type *5651*) then regulate the voltage to an extremely accurate value. The circuit and the construction of the power supply are aimed at keeping the conditions within the voltage reference tubes as constant as possible.

My observation of photomultiplier power supply performance indicated that not only does the output of the photomultiplier vary as the seventh power of voltage but also spurious pulses may be obtained by absorption of stray electromagnetic radiation (such as light, heat, radio waves, and gamma rays) by the voltage reference tubes. Any spurious or steady-state change in the internal energy of the gas of the voltage reference tube caused a change in its operating point. Therefore, the second set of tubes (type *5651*) was shielded from stray light and stray radio fields. They were placed in an oven to maintain constant thermal energy. An r-f filter was placed in the power supply to eliminate power-line noise. Since this experiment involved the use of radioactive materials, precautions were taken to insure that radiation energy was constant. When these precautions are taken, the output voltage can be maintained within 0.001 per cent. This degree of stability is necessary for wide-angle scintillation counters which are currently being used in nuclear research.

The purpose of this discussion has been twofold. First, a simple alternate power supply which is sufficiently accurate for photomultipliers as used in nuclear research has been described. Second, causes of reference voltage drift have been discussed. Adequate shielding for all types of radiation is necessary for very close regulation.

REFERENCE

1. PRECISION REGULATED HIGH VOLTAGE SUPPLIES, W. A. Higinbotham. *Review of Scientific Instruments*, New York, N. Y., vol. 22, June 1951, pp. 429–31.

W. G. Amey, F. H. Krantz, W. R. Clark, and A. J. Williams, Jr.: Mr. Jones' comments are very interesting. We used a 2-stage regulator circuit quite similar to Fig. 7 in our pH indicator which was completed in 1950. In this application a low voltage was needed so only two regulator tubes were required. The circuit performed very well for this application but the discontinuities of the first regulator stage were not sufficiently reduced by the *5651* tube to obtain a stability as good as 0.01 per cent. Several 105-volt and 150-volt regulator tubes were tried in the first stage but in each case good stability was possible only when the tubes operated in a region devoid of discontinuities. Since the locations of these discontinuities varied with time as well as from tube to tube, it was not possible to select a reliable operating region.

The high-voltage power supply described in the paper is free of these discontinuities and when operated with a constant load its regulation is better than 0.01 per cent. Since this performance was sufficient for our applications, no attempt was made to measure the regulation with greater precision. It should also be noted that the high-voltage power supply of the paper requires fewer tubes than the supply suggested by Mr. Jones.

Equations for Determining Current Distribution Among the Conductors of Busses Comprised of Double-Channel Conductors

CLIFFORD M. SIEGEL
ASSOCIATE MEMBER AIEE

THOMAS J. HIGGINS
MEMBER AIEE

A RECENT paper[1] advances equations enabling calculation of the inductance and short-circuit forces of busses comprised of double-channel conductors, as shown in Fig. 1 of that paper. These quantities are the essential parameters required for the rational design of this currently much-used bus. In an advertisement[2] it is stated that this type of construction is used in "11 out of the 16 largest turbine-generator" installations. The correctness and the usefulness of these equations in actual design are manifested by the valuable complementary discussion[1] of this recent paper given by O. R. Schurig, who compared values of reactance determined experimentally on a typical 3-phase bus comprised of 6-inch aluminum channels with values calculated from the equations of reference 1 and found good agreement among corresponding experimentally and analytically determined values.

Schurig's data encompasses measurement of the effective values of the currents in each of the six conductors comprising the bus: thus, 1,612 and 2,860, 2,755 and 2,400, 2,830 and 1,570 amperes for the six conductors, taken from left to right respectively. These values indicate that certain conductors carry substantially more current than others. Obviously, it is most desirable in regard to actual design that equations be available which afford ready calculation of the relative distribution of the total phase current between the two conductors of a

---

Paper 54-467, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Middle Eastern District Meeting, Reading, Pa., October 5–7, 1954. Manuscript submitted March 24, 1954; made available for printing July 26, 1954.

CLIFFORD M. SIEGEL is with the University of Virginia, Charlottesville, Va., and THOMAS J. HIGGINS is with the University of Wisconsin, Madison, Wis.

phase. This enables checking, during design, the possibility of overloading one or more conductors well above the permissible maximum average current density dictated by considerations of temperature rise during operation. The possibility of deriving such equations was indicated in the authors' closure[1] to Schurig's discussion. Accordingly, the prime purpose of the present paper is to advance these desired equations and to illustrate their use through calculation of the current distribution in two typical busses.

The present paper comprises:

1. Derivation, through pertinent use of the geometric mean distance (gmd) theory set out in reference 3 and certain results obtained in reference 1, of a basic equation enabling calculation of the fraction of the total phase current carried by each of the two conductors comprising a phase of the bus of Fig. 1.
2. Derivation of certain simple expressions enabling ready determination of the numerical values of the parameters of this equation for a specifically dimensioned bus.
3. Illustration of application of this equation through calculation therewith of the current division in one of the outer conductors of two typical busses, one of which has dimensions equal to that of the bus used in Schurig's experimental work.

The resulting calculated values for the latter bus prove to be in good agreement with the corresponding values stemming from Schurig's aforementioned experimental data, which agreement manifests both the correctness of the general analysis and its usefulness in actual design.

Derivation of the Basic Equation

For the bus of Fig. 1 the voltage drop along each of the two channels comprising a conductor must be equal. If the resistive drop is much smaller than the reactive drop, and this is actually the case in practice, the two reactive voltage drops must be essentially equal. In such a case, on equating the two, $2\pi f L_{k1} I_{k1} = 2\pi f L_{k2} I_{k2}$; thus

$$L_{k1} I_{k1} = L_{k2} I_{k2} \tag{1}$$

where the subscripts $k1$ and $k2$ designate the two channels comprising the kth conductor of Fig. 1; L_{k1} and L_{k2} are the inductances associated with the two channels, and I_{k1} and I_{k2} the corresponding currents in the two channels.

From equation 1 of reference 1

$$L_{k1} = -2\left(\sum_{\substack{i=1\\ i\neq k}}^{r} (I_i/I_{k1}) \log D_{i,k1} + (I_{k2}/I_{k1}) \log D_{k1,k2} + (I_{k1}/I_{k1}) \log D_{k1,k1}\right) \tag{2}$$

and

$$L_{k2} = -2\left(\sum_{\substack{i=1\\ i\neq k}}^{r} (I_i/I_{k2}) \log D_{i,k2} + (I_{k2}/I_{k2}) \log D_{k2,k2} + (I_{k1}/I_{k2}) \log D_{k2,k1}\right) \tag{3}$$

where

$D_{i,k1}$ = gmd between the ith conductor and the left channel of the kth conductor

$D_{i,k2}$ = gmd between the ith conductor and the right channel of the kth conductor

$D_{k1,k2} = D_{k2,k1}$ = gmd between the two channels of the kth conductor

$D_{k1,k1} = D_{k2,k2}$ = self-gmds of the left and right channels of the kth conductor respectively

The dimensions of all quantities are to be expressed in the centimeter-gram-second system of electromagnetic units, which are particularly appropriate to calculation with gmd theory. Substituting equations 2 and 3 in equation 1, collecting terms and simplifying, yields

$$\sum_{\substack{i=1\\ i\neq k}}^{r} I_i \log (D_{i,k1}/D_{i,k2}) + I_{k2} \log (D_{k1,k2}/D_{k2,k2}) + I_{k1} \log (D_{k1,k1}/D_{k2,k1}) = 0 \tag{4}$$

Inasmuch as $D_{k1,k1} = D_{k2,k2}$, equation 4 reduces to

$$\sum_{\substack{i=1\\ i\neq k}}^{r} I_i \log (D_{i,k1}/D_{i,k2}) + (I_{k2} - I_{k1}) \log (D_{k1,k2}/D_{k1,k1}) = 0 \tag{5}$$

Finally, division of each side of equation 5 by I_k and solvng as evident yields the following expression for the difference of the currents in the two channels comprising the conductor, expressed as a fraction of the total conductor current

$$(I_{k1} - I_{k2})/I_k = \left(\sum_{\substack{i=1\\ i\neq k}}^{r} (I_i/I_k) \times \log (D_{i,k1}/D_{i,k2})\right) / \log (D_{k1,k2}/D_{k1,k1}) \tag{6}$$

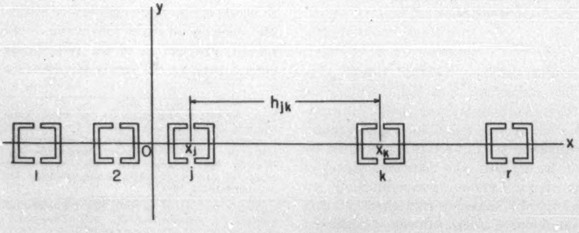


Fig. 1. Schematic of coplanar r-phase bus

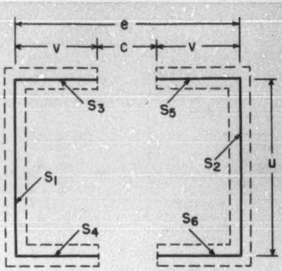


Fig. 2. Line segment approximation of conductor cross section

Calculation of GMD

As stated in reference 1 and confirmed therein by calculation, most channels used in practice are so thin that excellent approximations for the actual values of the gmd indicated in equation 6 result if the actual cross sections of the channels are replaced by line segment cross sections having dimensions equal to the corresponding mean dimensions of the actual bus; see Fig. 2. In this case, a straightforward though lengthy calculation, carried out as detailed in reference 3, yields

$$\log (D_{i,k1}/D_{i,k2}) \cong [2e(i-k)/(2v+u)|i-k|] \tan^{-1} (u/2h) \tag{7}$$

and

$$\begin{aligned}\log (D_{k1,k2}/D_{k1,k1}) = {} & [2e^2 \log e - 2v^2 \log v - 3u^2 \log u + c^2 \log c + 2ue \log (e^2+u^2) - \\ & ((c+v)^2 + 2u(c+v) - u^2) \log ((c+v)^2+u^2) - 2(c+v)^2 \log (c+v) + (c^2-u^2) \log \\ & (c^2+u^2)^{1/2} - (v^2+2uv-u^2) \log (u^2+v^2) + 2(e^2-u^2) \times \tan^{-1} (u/e) + 2((c+v)^2 - u^2 + \\ & 2u(c+v)) \times \tan^{-1} ((c+v)/u) + 2uc \tan^{-1} (c/u) - 2(u^2+2uv-v^2) \tan^{-1} (v/u) + \\ & \pi(u^2+eu-v^2-(c+v)^2)]/(u+2v)^2\end{aligned} \tag{8}$$

A further approximation for the second of these two functions results from obviating the gap between the two channels, thus letting c approach zero, while simultaneously increasing v so that $c+2v$ remains constant and equal to e; see Fig. 3. In this case, substitution of $c=0$ and $v=e/2$ in equation 8 yields

$$\log(D_{k1,k2}/D_{k1,k1}) = [2e^2\log e - e^2\log(e/2) - 4u^2\log u + 2ue\log(e^2+u^2) - 2(ue-u^2+(e/2)^2)\log(u^2+(e/2)^2) + 2(e^2-u^2)\tan^{-1}(u/e) + (4u^2+4ue-e^2)\tan^{-1}(2u/e) - \pi u(e+u)]/(u+e)^2 \quad (9)$$

An alternative form of equation 9 which is more convenient for computation

$$\log(D_{k1,k2}/D_{k1,k1}) = (-\pi U(U+1) + 2U\log(1+U^2) - (0.5-2U^2+2U)\log(0.25+U^2) + 2(1-U^2)\times\tan^{-1}U - (1-4U^2-4U)\times\tan^{-1}2U - 4U^2\log U + \log 2)/(U+1)^2 \quad (10)$$

where U is a dimensionless parameter defined by $U=u/e$.

A graph of equation 10 is shown in Fig. 4, plotted from the values tabulated in Table I. This graph and the simple form of the right-hand member of equation 7 enables easy computation of the terms in the right-hand member of equation 6, and of the division of current between the two channels of a conductor. In carrying out this computation, it is to be noted that the dimensions of the cross section of the bus can be expressed in any convenient unit of length, and the currents in the conductors can be expressed in any convenient unit of current; for equation 6 is couched in terms of the ratios of currents and the ratios of gmd and, therefore, is expressed in terms of dimensionless quantities.

An Illustrative Example

To show the use of the equations derived in this paper, the solution of the following problem is illustrated. Consider a 3-conductor 3-phase bus, oriented as shown in Fig. 1, comprised of standard copper channels and having the dimensions: web $a=4$ inches; flange width $f=1.75$ inches; thickness $t=0.2$ inch; separation between channels $c=1.25$ inches; spacing between centers of conductors $h=28$ inches. Assume that the bus carries balanced 3-phase currents; though in view of the unbalanced coplanary geometry of the bus of Fig. 1, the phase currents will, of course, not be exactly balanced. However, calculation and experimental measurement of the actual total current in each of the three phases of typical busses yield that the unbalance is very slight; therefore, the assumption of balanced phase currents is sufficiently accurate for present purposes.

Thus

$$I_1 = I_2\epsilon^{-j120\text{ degrees}} = I_3\epsilon^{j120\text{ degrees}}$$

Find the current division between the two channels of the outer conductor, carrying current I_1.

Simple calculation with the given dimensions yields: $u=3.8$ inches; $v=1.65$ inches; $e=4.55$ inches; $h_{12}=h_{23}=28$ inches; $h_{13}=56$ inches; number of conductors $r=3$; conductor of interest $k=1$. For this problem equation 6 becomes

$$(I_{11}-I_{12})/I_1 = \left(\sum_{i=2}^{3}(I_i/I_1)\times \log(D_{i,11}/D_{i,12})\right)/\log(D_{11,12}/D_{11,11}) \quad (11)$$

Now, $I_2/I_1 \cong \epsilon^{j120\text{ degrees}}$ and $I_3/I_1 \cong \epsilon^{-j120\text{ degrees}}$, whence on using equation 7

$$(I_{11}-I_{12})/I_1 = \frac{(\epsilon^{j120\text{ degrees}}\tan^{-1}(u/2h_{12}) + \epsilon^{-j120\text{ degrees}}\tan^{-1}(u/2h_{13}))(2e/(2v+u))}{\log(D_{11,12}/D_{11,11})} \quad (12)$$

Table I. Data for Fig. 4

u/e	log ($D_{k1,k2}/D_{k1,k1}$)
0	1.387
0.25	1.076
0.50	0.983
1.0	0.861
2.0	0.677
3.0	0.555
4.0	0.480

Calculation of $\log(D_{11,12}/D_{11,11})$ from equation 10 where $U=u/e=0.834$ yields a value of 0.89 in agreement with the value to be obtained directly from the curve of Fig. 4. Substituting this value and the others indicated in equation 12 and effecting the necessary computation gives

$$(I_{11}-I_{12})/I_1 = -0.0735+j0.0425 \quad (13)$$

Representing this ratio by the convenient symbol R and recalling that $I_{11}+I_{12}=I_1$, it is found that $I_{11}/I_{12}=(1+R)/(1-R)$; thus, by use of equation 13

$$I_{11}/I_{12} = (0.9265+j0.0425)/(1.0735-j0.0425) = 0.86\epsilon^{j4.9\text{ degrees}} \quad (14)$$

Finally, from equation 14 the desired ratio $|I_{11}/I_{12}|$ of the effective values of the two currents is found to be 0.86. Thus, the outer channel carries somewhat less current than the inner channel. This is in general accord with the distribution found experimentally on a bus of the same general structure by Schurig, as mentioned in the outset of the paper.

Comparison of Calculated and Experimental Values

In conclusion, and substantiative of the correctness of the analysis in general, the calculation of the current division in the outer conductor of the bus used in Schurig's investigations is illustrated. This bus comprises a 3-phase aluminum-channel structure, as shown in Fig. 1: web $a=6$ inches; flange width $f=2.157$ inches; web thickness $t=0.2$ inch; cross

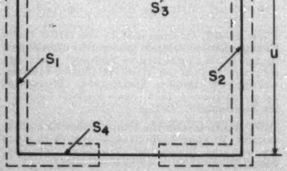


Fig. 3 (left). Line segment approximation of conductor cross section, neglecting gap

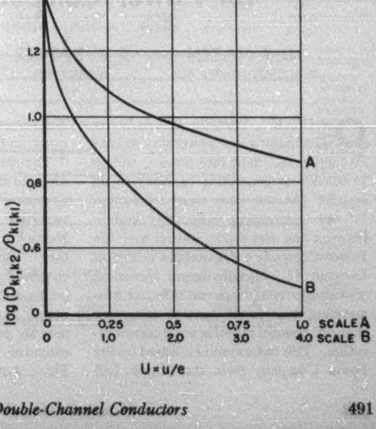


Fig. 4 (right). Graph of equation 10

Table II. Comparison of Calculated and Measured Values

	I_{11}/I_{12}		I_{21}/I_{22}		I_{31}/I_{32}		I_{32}/I_{31}	
Calculated	0.613	−17°	1.0	−56.7°	1.63	−17°	0.613	17°
Measured	0.563	−22°	1.14	−63°	1.8	−8°	0.55	8°

section of channel assembly forms a hollow "square" $c+2f=a=6$ inches; spacing between adjacent centers of conductors $h=12$ inches; $I_1 \cong I_2 \epsilon^{j120 \text{degrees}} \cong I_3 \epsilon^{-j120 \text{degrees}}$.

Simple calculation with these values yields: $v=1.939$ inches; $e=u=5.563$ inches; $c=1.685$ inches; $h_{12}=h_{23}$ 12 inches and $h_{13}=24$ inches; $r=3$; conductor of interest $k=1$. Substituting in equation 11 and using equation 7 yields an expression identical with equation 12, except for the interchange of +120 degrees and −120 degrees resulting from the opposite phase sequences of the two problems. Effecting computation with the present values as detailed between equations 12 and 13 gives

$$R=(I_{11}-I_{12})/I_1=-0.246-j0.138 \qquad (15)$$

Finally, from $I_{11}/I_{12}=(1+R)/(1-R)$ expressed in polar form, yields

$$I_{11}/I_{12}=0.613\epsilon^{-j17 \text{ degrees}} \qquad (16)$$

Thus, as in the first illustrative problem, the outer channel carries less current than the inner. It should be noted that the current in the former lags the latter, whereas in virtue of the reversed phase sequence the opposite is true in the first illustrative problem, as shown in the complex number ratio of equation 14.

Confirmatively, the corresponding ratio as calculated from the experimental values measured by Schurig is

$$I_{11}/I_{12}=0.563\epsilon^{-j22 \text{ degrees}} \qquad (17)$$

Considering the geometrical approximations, the neglect of skin-effect, which has a finite though well-established[4,5] small effect on the conductor inductances, the assumption of balanced phase currents, the neglect of phase resistance, and the usual degree of error to be expected in the experimental measurement of the large (slightly differing in phase) currents flowing in the two channels of a conductor, the calculated ratio of 0.613 and the experimentally determined value of 0.563 may be considered in fair agreement. Further, substantially the same degree of agreement between calculated and experimentally determined values is to be found for the second and third conductors, as shown in Table II. This confirms the possibility of a statement made in the authors' closure to their previous paper,[1] that "in fact, if desired, these two currents (in the channels of a conductor) could be calculated rather closely by use of the theory given." Hence, the theory given in this paper can be used, as illustrated in the foregoing two examples, to calculate the distribution of currents among the two channels of a conductor with sufficient accuracy for purposes of bus design.

In conclusion, it should be noted that neglect of the gap in a conductor, as so indicated in Fig. 3, enables approximate calculation of the distribution of current in the two halves of a conductor of a 3-phase coplanar bus comprised of square or rectangular hollow tubular conductors—a geometry also much used in practice. Such determination is, as with double channel conductors, of interest in regard to design as it enables a sufficiently satisfactory estimate of the average current density over the most heavily loaded portions of the two outer conductors of the bus.

References

1. Equations for the Inductance and Short-Circuit Forces of Busses Comprised of Double-Channel Conductors, **C. M. Siegel, T. J. Higgins.** *AIEE Transactions*, vol. 71, pt. III, 1952, pp. 522–31.

2. General Electric advertisement. *Electrical Engineering*, vol. 69, 1950, p. 43A.

3. Formulas for the Calculation of the Inductance of Linear Conductors of Structural Shape, **Thomas James Higgins.** *AIEE Transactions (Electrical Engineering)*, vol. 62, Feb. 1943, pp. 53–57.

4. Formulas for the Inductance of Coaxial Busses Comprised of Square Tubular Conductors, **Henry Peter Messinger, Thomas James Higgins.** *AIEE Transactions (Electrical Engineering)*, vol. 65, June 1946, pp. 328–36.

5. The Inductance of Linear Conductors of Rectangular Section, **A. H. M. Arnold.** *Journal*, Institution of Electrical Engineers, London, England, vol. 70, 1931–32, pp. 579–86.

◆

No Discussion

A Private Microwave Radio System for Power Company Use

D. F. HAZEN
ASSOCIATE MEMBER AIEE

J. W. DANSER
MEMBER AIEE

G. S. ZILIS
ASSOCIATE MEMBER AIEE

ON JUNE 1, 1953, the Illinois Power Company's new power station near Hennepin, Ill., went into service, adding 75,000 kw of generating capacity to its system. At the same time a 6-station 112-mile microwave radio relay system between the Hennepin Station and the Power Company's dispatching center in Decatur, Ill., officially began operation, providing over 1,300 circuit miles for telemetering and remote control, and more than 400 channel miles for voice communication. This radio system is leased by the Power Company from the Illinois Bell Telephone Company which engineered, owns, and maintains it.

The channels provided by this system are used to facilitate control of the Power Company's system and co-ordinate its operation with neighboring utilities in the North Central part of its operating territory. Sixteen telemeter and control channels are provided between various points either directly or via power-line carrier extensions. Four talking channels are provided and some of these are extended by means of cable facilities. Fig. 1 shows the geographical location of the system and its extensions, and Fig. 2 is a diagram of the voice-frequency channel arrangement and the circuits provided.

It is the purpose of this paper to relate the joint experience of a power company and a telephone company in engineering and maintaining a microwave system and also to describe the salient features of this system. Some of the problems encountered with emphasis on the propagation difficulties are discussed. The provision of telemetering channels over this system is treated in a separate section in which the effects of microwave operation on the

---

Paper 54-407, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 14, 1954; made available for printing September 27, 1954.

D. F. Hazen is with the Illinois Power Company, Decatur, Ill., and J. W. Danser and G. S. Zilis are with the Illinois Bell Telephone Company, Chicago, Ill.

metering and load control system are also discussed.

Genesis of System

The Power Company foresaw the need for a large group of communications channels north from its dispatching center at Decatur during the planning for the Hennepin station and the Company's participation in the Illinois-Missouri Electric Power Pool which involved commitments to the Atomic Energy Commission. It was decided that a microwave system was the economical answer to the problem. Specifications for bids for the installation and maintenance of the proposed system were issued to the manufacturers and suppliers. The Telephone Company entered its bid for this system on the basis of the use of microwave because it was felt that it was the most economical facility consistent with the requirements in this case. The Telephone Company was the successful bidder, agreeing to lease the system to be used solely to provide channels for the Power Company. The Power Company would be the radio licensee and operation would be on radio frequencies allocated for industrial radio services.

System Engineering

Features which were deemed desirable for the radio equipment to be used for this system were: reliability, simplicity, and a maximum capacity of 15 to 30 voice frequency channels. The light route nature of this system would not justify the use of normal Bell System radio relay equipment as used in heavy route long-haul service. The 6,800-megacycle equipment with complete radio-frequency (r-f) stand-by and 24 channel-type frequency-modulated frequency-division multiplex manufactured by the Motorola Company was selected for this system. A feature of this equipment which is pertinent to the foregoing requirements is worthy of mention. The r-f portion of the microwave transmitter contains only one electronic device. This is a relatively trouble-free klystron oscillator which is modulated by a simple voltage amplifier.

The Power Company provides the radio station sites, except at Bloomington and Benson. The Decatur, Bloomington, Oglesby, and Hennepin stations are those involving telemetering points of utilization and their locations were chosen to be within 1/4 mile of these points. Buried cable or cable in conduit is used to extend the telemetering circuits from the microwave equipment.

The Benson station location was the

Fig. 1 (right). Route of Decatur - Hennepin microwave radio system

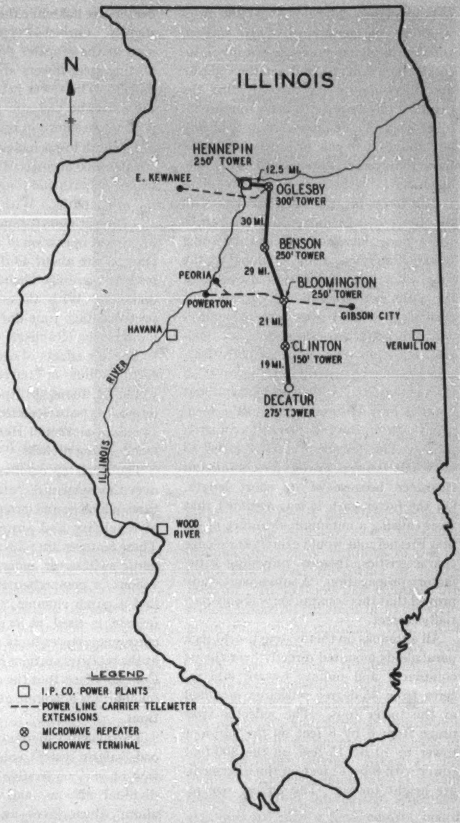
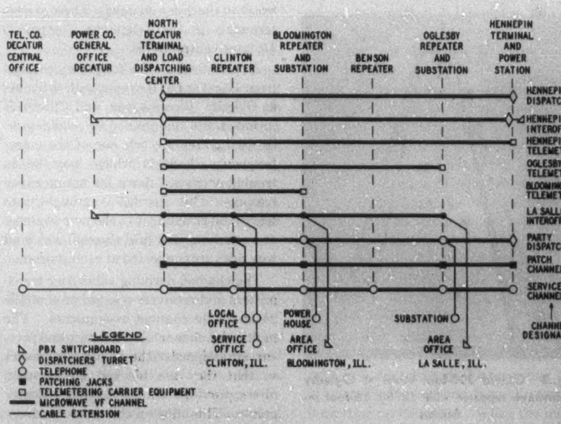

N
ILLINOIS
HENNEPIN
250' TOWER
12.5 MI.
E. KEWANEE
OGLESBY
300' TOWER
30 MI.
BENSON
250' TOWER
29 MI.
PEORIA
BLOOMINGTON
250' TOWER
POWERTON
GIBSON CITY
21 MI.
HAVANA
CLINTON
150' TOWER
VERMILION
19 MI.
DECATUR
275' TOWER
ILLINOIS RIVER
WOOD RIVER
LEGEND
I. P. CO. POWER PLANTS
POWER LINE CARRIER TELEMETER EXTENSIONS
MICROWAVE REPEATER
MICROWAVE TERMINAL

Fig. 2. (below) Voice-frequency channel arrangement


TEL. CO. DECATUR CENTRAL OFFICE
POWER CO. GENERAL OFFICE DECATUR
NORTH DECATUR TERMINAL AND LOAD DISPATCHING CENTER
CLINTON REPEATER
BLOOMINGTON REPEATER AND SUBSTATION
BENSON REPEATER
OGLESBY REPEATER AND SUBSTATION
HENNEPIN TERMINAL AND POWER STATION
HENNEPIN DISPATCH
HENNEPIN INTEROFFICE
HENNEPIN TELEMETER
OGLESBY TELEMETER
BLOOMINGTON TELEMETER
LA SALLE INTEROFFICE
PARTY DISPATCH
PATCH CHANNEL
SERVICE CHANNEL
CHANNEL DESIGNATION
LOCAL OFFICE
SERVICE OFFICE
CLINTON, ILL.
POWER HOUSE
AREA OFFICE
BLOOMINGTON, ILL.
SUBSTATION
AREA OFFICE
LA SALLE, ILL.
LEGEND
PBX SWITCHBOARD
DISPATCHERS TURRET
TELEPHONE
PATCHING JACKS
TELEMETERING CARRIER EQUIPMENT
MICROWAVE VF CHANNEL
CABLE EXTENSION

only one mainly determined by radio propagation considerations. This station, which is simply a microwave repeater, is located about halfway between the Oglesby and Bloomington stations, dividing the distance into two paths each roughly 30 miles long. There did not seem to be any advantage, economic or otherwise, to use more than one repeater over this 60-mile distance.

The tower heights were decided upon the basis of obtaining path clearances which were considered sufficient according to the best information available at the time. Earth profiles and field checks of the paths were made noting obstructions such as trees and man-made structures whose heights are not shown on topographic maps. The minimum path clearances assuming true earth radius over all obstructions including trees, were about 1.1 first Fresnel zone except on the Hennepin-Oglesby and Benson-Bloomington paths. The former was engineered to have only 0.5 first Fresnel zone minimum clearance because of its short length. On the latter path it was assumed that trees causing a minimum clearance of 0.6 first Fresnel zone would exhibit knife-edge characteristics, thereby providing satisfactory propagation. A subsequent study proved that this assumption was only partially correct.

All antennas on this system use 40-inch parabaloids mounted directly over the r-f equipment and radiate toward and receive from 45-degree reflectors mounted at the tower tops. The reflector sizes range from 6 by 8 feet on the 150-foot tower to 10 by 15 feet on the 300-foot tower with 8 by 12 feet on the intermediate height towers. The largest size reflectors are flat while the others are curved slightly. Guyed H-type steel towers are used at the repeater points and self-supporting steel towers are used at the terminals. The tower heights are shown in Fig. 1. The Oglesby tower and its reflectors are shown in Fig. 3.

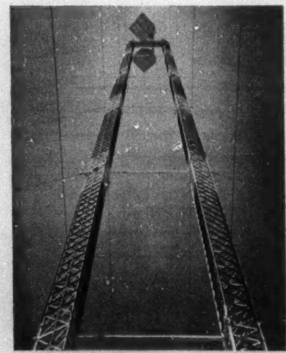


Fig. 3 Guyed 300-foot tower at Oglesby microwave repeater with 10 by 15-foot reflectors

The path losses including antenna gains are about 60 decibels (db) on the Oglesby-Hennepin path and roughly 65 to 68 db on the other paths. With a maximum permissible path loss for successful telemetering circuit operation of 102 db, the fading margins are about 43 db for the shortest path and average about 36 db for the remainder. With these margins the expected outage time due to multiple path fading only (Rayleigh distribution, neglecting the effects of changes in apparent earth radius or terrestrial reflections) would be about 0.12 per cent on 1-way impulse-type telemetering circuits between Decatur and Hennepin during the worst fading months.

The type of communications carried over this system dictated the use of features which would provide a high degree of reliability and stormproof operation. These features are r-f stand-by with automatic switchover, emergency power provisions, a comprehensive alarm system, and a patch channel. The r-f stand-by feature is used at every station while emergency power units are provided only at the repeater stations because the Power Company feels that the commercial power reliability is adequate at the terminal stations.

Two alarm systems are employed, one a coded alarm system which gives an indication of various trouble conditions at individual stations, and the other a loop alarm which gives an indication of a break in the video channel. These alarms terminate in the Telephone Company's Decatur central office.

A spare 2-way voice-frequency channel from Decatur to Hennepin with dropouts at Oglesby, Bloomington, and Clinton is included and maintained in readiness to be used to replace any one of the voice-frequency channels which may be in trouble or turned down for maintenance reasons. This channel is brought into service by patching it to the wire facilities at a patching jack field through which all wire lines are connected at each station.

The channel deriving subcarrier transmitters and receivers operate on 9 of the 24 available channel assignments. The multipoint channels, both future and present, are assigned to the higher frequencies so that they are blocked by low-pass filters provided at the desired repeater points. This allows a channel to be dropped and reinserted at repeater stations. The LaSalle interoffice, the party dispatch, and the service and patch channels, as shown in Fig. 2, are the multipoint lines of this system.

Each talking channel including the party lines operates on a full duplex basis and is set up to provide a loss of 10 db between any two switchboards or turrets and 16 db between any two telephones directly connected to the system. These telephones are of the 4-wire type. The voice channels including the cable extensions are mainly 4-wire with talk-back bridges used where bridged stations are involved. At the switchboards and turrets each voice channel is converted to 2-wire operation so that the channel is available to any telephone terminating on these units.

The talking circuits are almost completely protected against commercial power failures. All telephones directly connected to the system are of the local talking battery type and use push-button signalling with the source of signalling power located at the microwave station where reliable power is available. Each switchboard and turret is powered by batteries or is connected to a telephone central office battery supply.

Microwave Equipment Description

Two types of station installations are used. One type has all equipment except the antenna housed in a small metal building located at the base of the antenna tower as shown in Fig. 4. In the other type the r-f units are mounted in an enclosed weatherproof cabinet located adjacent to the tower and remote control equipment; power supplies and multiplex equipment are located in a near-by building.

The microwave transmitter consists of a klystron oscillator which is frequency-modulated by varying its repeller voltage with the output of a 2-tube 1-megacycle video amplifier. The receiver is of the

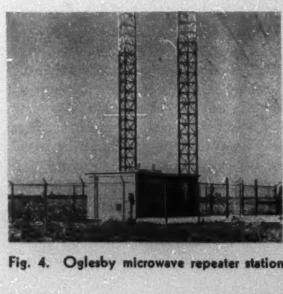


Fig. 4. Oglesby microwave repeater station

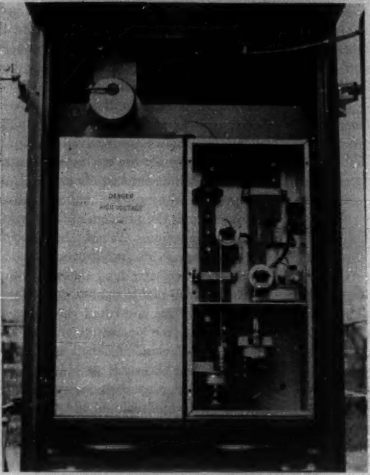


Fig. 5 (left). Main and stand-by r-f units of remotely controlled type showing wave-guide plumbing and klystrons of stand-by unit

Fig. 6 (right). Technician adjusting remote control type r-f unit

superheterodyne type using a crystal mixer, klystron local oscillator, a 75-megacycle intermediate-frequency strip, and an automatic-frequency-control chassis. The transmitter and receiver with their associated wave-guide plumbing are contained in an insulated box as shown in Figs. 5 and 6.

The multiplex equipment is of the frequency-division type and consists of subcarrier frequency-modulated transmitters and receivers, voice terminal units, and the associated power supplies. At intermediate points, the party lines are dropped and reinserted by the use of heterodyne-type multiplex units which are equivalent to 3-way 4-wire bridges.

Each microwave station has a coded alarm transmitter connected to the video bus transmitting toward Decatur. When an alarm condition exists a cam arrangement keys the transmitter output with a code indicating a particular emergency. At Decatur the alarm signals received at the microwave station are sent to the Telephone Company's central office over an exchange cable pair. Here these signals are applied to an alarm detector which sounds the alarm. R-f stand-by and emergency power operation and tower light failures are indicated by this alarm system.

System Maintenance

The system is maintained by routine and trouble visits to the stations. Routine visits are made at intervals not exceeding 2 weeks and the various routine checks are made at regular intervals varying from 2 weeks to 6 months.

Centralized maintenance control has been established at the Decatur central office which is attended 24 hours a day. A toll testman assigned to the microwave system is informed of trouble conditions either from the alarm apparatus or from the Decatur repair service operator who obtains trouble reports by telephone calls from Power Company personnel. During trouble periods this testman dispatches maintenance technicians and communicates with them at the various microwave stations by means of the service channel.

Maintenance of the system is divided into three sections; each includes two adjacent stations. There are at least two maintenance men available in each section, one assigned for routine checks and either or both on call in case of system trouble. The maintenance technicians have had formal training in the operation and maintenance of the microwave equipment being used. In addition, these men maintain at least two of the following, microwave and wire television links, mobile radio service, telephone repeater and carrier equipment, and teletypewriter apparatus.

Telemetering

Operation of the Hennepin Generating Station together with the establishment of intersystem ties between the Illinois Power Company and neighboring company networks at Oglesby, Kewanee, Gibson City, Peoria, and Powerton make necessary the provision of 12 telemeter channels from these locations to the Illinois Power load dispatcher at North Decatur. In addition four channels are required from North Decatur to the Hennepin Station to transmit "raise" and "lower" governor control signals and "area" and "station" requirement indications to the generating station.

At the North Decatur dispatcher's office megawatt and megavar indications are received from Hennepin and each of the interchange points. Megawatt information, as well as indications of system frequency, power interchange schedules, etc., are fed into a load control board which automatically integrates the information and sends appropriate "raise" or "lower" signals to the Hennepin Generating Station.

Fig. 7 indicates the arrangement of these 16 channels on the microwave system. They are provided by audio-frequency carrier equipment[1] operating over three voice-frequency channels as shown in Fig. 2. Leeds and Northrup terminal equipment of both the impulse-duration and continuous types is used. Impulse-duration type equipment is used between Decatur and the Hennepin Station. The continuous type is used for transmission of megawatt information from the interchange points, while the impulse-duration type transmits the megavar information.

Interruption of a telemeter channel results in a departure of the recording pen of the telemeter receiver from the correct indication on the chart. When the inter-

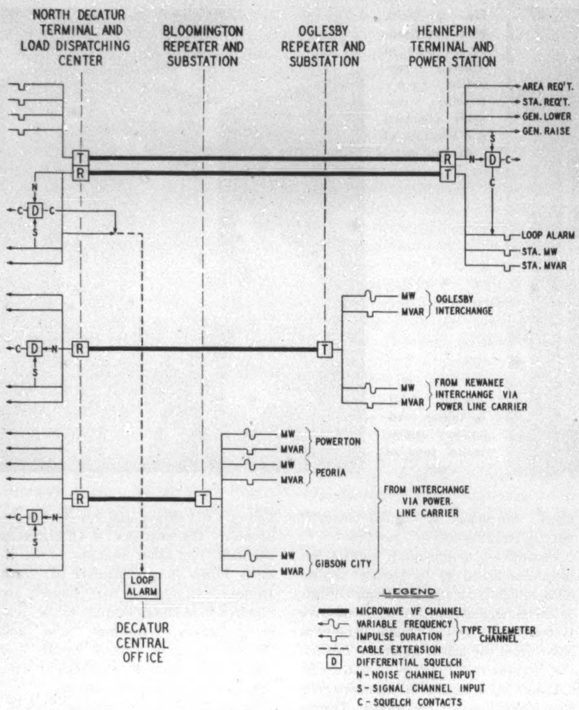

NORTH DECATUR TERMINAL AND LOAD DISPATCHING CENTER
BLOOMINGTON REPEATER AND SUBSTATION
OGLESBY REPEATER AND SUBSTATION
HENNEPIN TERMINAL AND POWER STATION
AREA REQ'T.
STA. REQ'T.
GEN. LOWER
GEN. RAISE
T
R
D
N
S
C
LOOP ALARM
STA. MW
STA. MVAR
MW
MVAR
OGLESBY INTERCHANGE
FROM KEWANEE INTERCHANGE VIA POWER LINE CARRIER
POWERTON
PEORIA
GIBSON CITY
FROM INTERCHANGE VIA POWER LINE CARRIER
LOOP ALARM
DECATUR CENTRAL OFFICE
LEGEND
MICROWAVE VF CHANNEL
VARIABLE FREQUENCY
IMPULSE DURATION
TYPE TELEMETER CHANNEL
CABLE EXTENSION
D DIFFERENTIAL SQUELCH
N - NOISE CHANNEL INPUT
S - SIGNAL CHANNEL INPUT
C - SQUELCH CONTACTS

Fig. 7. Telemeter channel arrangement

ruption is due to fading, as might be the case in channel operation over microwave, the attendant noise in the receiver is usually high enough to drive the recording pen erratically across the chart. This erratic operation often results in tearing of the chart or breaking of the pen. Further, if the interrupted channel is one which provides information to the load control board, departure of the recorder indication to positions outside preset limits trips the board from automatic to manual operation. To return to automatic control the operator must remove the channel from the board and change a miscellaneous input setting by an amount equal to the last correct indication received over the interrupted channel.

While fading outages can be minimized by provision of adequate fading margins in the microwave system, it does not appear economically feasible at present to reduce them to the point where disturbances in the control board operation will be of no consequence. Experience indicates that fading occurs to interrupt the telemetering signals continuously for only a short time, usually in the order of seconds and rarely for more than a few minutes. While intermittent channel interruptions may continue for several hours during periods of severe fading, the telemetering signal is re-established frequently enough to provide a useful indication of the metered quantity.

To avoid the difficulties discussed in the foregoing, four differential-type squelch units have been incorporated with the audio-frequency carrier equipment. These are shown in Fig. 7. The operation of these units has been discussed elsewhere.[1] Squelch arrangements provided with the voice-frequency channels over which the audio carrier channels operate have been disabled, since they cannot be adjusted for the low signal-to-noise ratio conditions under which satisfactory telemeter signal reception can be maintained.

The four differential squelch units provide contact arrangements for use with auxiliary equipment added to the telemeter recorders to de-energize the pen motors during squelch operation and to re-energize them after the channels have been unsquelched and one or two complete telemeter signal cycles received. This arrangement permits the telemeter recorder pen to remain at the last correct indicating position during the channel outage and overcomes the problem of torn charts and broken pens. While the squelch operation brings in an alarm at the dispatcher's office, the load control board need not be tripped to manual control immediately. Wide and rapid variations of conditions in the operation of a power system are not usual and the loss of information during the short periods of complete fading outage, which would normally be expected, should not be objectionable. Studies are now in progress to determine how long the control board should be permitted to continue on automatic operation after the loss of telemetered information.

As mentioned previously, a "loop alarm" arrangement has been established by the use of an additional telemeter channel and auxiliary contacts on two of the squelch units. This alarm supplements the coded alarm system and indicates a break in the video channel in either direction along the microwave route, a condition which the coded alarm system fails to indicate. Such a break could result from failure of emergency power, automatic switchover, or r-f stand-by equipment. Interruption of the video channel toward Hennepin operates the differential squelch at Hennepin. Indication of this operation is transmitted back to Decatur over the additional telemeter channel. Similarly interruption of the video channel in the Decatur direction operates the differential squelch at Decatur associated with the Hennepin-Decatur telemeter channels. Indication of the operation of either squelch unit is provided to the Decatur central office over cable facilities to initiate an alarm when a continuous squelch condition persists longer than 2 minutes.

System Performance

The performance of the system generally has been satisfactory. While some outages occurred because of radio propagation difficulties, the availability of the system for telemetering over its full length has been in the order of 99.6 per cent. Fading affects telemetering channel performance in a positive manner, but does not cause serious trouble on the voice circuits due to the fact that a great part of the fading occurs during hours when these circuits are not in use. The r-f equipment has been particularly reliable because of the stand-by feature.

To obtain data on system performance, the Telephone Company engineering department conducted a 4-month study of the continuity of service shortly after the system was put into service. This consisted of an analysis of continuous recordings of signal strength on several of the r-f paths and of the Power Company's telemetering charts which at the time gave an indication of fading outages. The results indicated that the video channel was continuous and in suitable condition for proper telemetering transmission from Hennepin to Decatur for about 99.6 per cent of the study time. This represents an average outage time of about 5 minutes per day. Since this study was not completely comprehensive the following breakdown of the outages is not highly accurate. The conditions responsible and the percentage outage time during the study (7-20-53 to 11-20-53) were as follows.

Radio signal fading, 0.2 per cent
Equipment failures, 0.05 per cent
Emergency power failures, 0.04 per cent
Maintenance and engineering tests, 0.1 per cent

A breakdown of the fading outages in this study indicated that the outages on the Benson-Bloomington and the Oglesby-Benson paths were in excess of the values expected from pure multipath fading. The outage time on the Oglesby-Benson path was over twice this expectation and was about half the system fading outage time. The Benson-Bloomington path fading outages were about 50 per cent over the multipath expectation.

In general the seasonal and diurnal variations of fading observed in the study were similar to those observed in propagation tests by Durkee,[2] a seasonal maximum occurring during the summer and daily minimum fading around the noon hour. Very few outages due to fading occurred during the colder parts of the year.

An examination of the character of the signal strength recordings in the foregoing indicated that the fading was of well-known types which have been described elsewhere.[3] These types occurred individually or in combination. There was the multiple path type which consisted of relatively rapid variations of signal strength with the total time of excursions from maximum to minimum in the order of seconds and minutes. In addition there were large depressions of the average signal lasting as long as several hours in the early morning during the late summer and early autumn. These were attributed to inverse bending which resulted from conditions associated with ground fog of the radiation type. The recording in Fig. 8 shows a severe depression of the signal strength between 3 a.m. and 6 a.m. believed to have been caused by inverse bending.

The fading outages were intermittent and usually short, but a number of them were of 1 or 2 minutes' duration. There were a few that lasted as long as 5 minutes but outages longer than this were rare. Most of the outages over 1 minute long occurred during periods when inverse bending was suspected.

During the course of the study just mentioned, due to the fact that some individual path outage times were higher than expected, a detailed recheck of the clearances on all but the Hennepin-Oglesby path was made. On the two longest "hops" the exact location of the paths over the critical portions of their lengths was determined by surveying methods using the microwave tower beacon lights as base points.

On the Benson-Bloomington path this survey revealed that removal of a group of eight trees, some of which were 65 feet high, would increase the minimum clearance on this path by about 0.5 first Fresnel zone. The trees were removed leaving this path with a minimum vertical clearance of 1.1 first Fresnel zone assuming true earth radius. A study of the signal strength recordings indicated that during periods when inverse bending was suspected these trees caused a grazing condition which resulted in an average signal depression of 10 db or more. Since the trees were removed late in October 1953, the expected results will not be apparent until summer fading resumes.

The survey also revealed that assuming an effective earth radius K equal to 2/3 the Oglesby-Benson path has a number of points of minimum clearance in the order of 0.2 to 0.3 first Fresnel zone extending over a distance of about 10 miles. It is believed that this is a marginal condition and is the cause of deep fading during periods of inverse bending. The possibility of a moderate reflection from the earth on this path was also indicated. A discussion of some of the theoretical concepts involved in this paragraph is contained in a paper by R. D. Campbell.[4]

A new series of continuous recordings using five Esterline Angus recorders was started on February 1, 1954. These will be used to obtain further data on system performance. The signal level is being recorded on all but the shortest path. Differential squelch relay operations on the Hennepin to Decatur and Bloomington to Decatur telemetering channels are also being recorded.

The performance of the klystron oscillators has been good; during the first 6 months of operation there were only three "in service" failures and these were adequately cared for by switchover to the stand-by equipment. There have been a

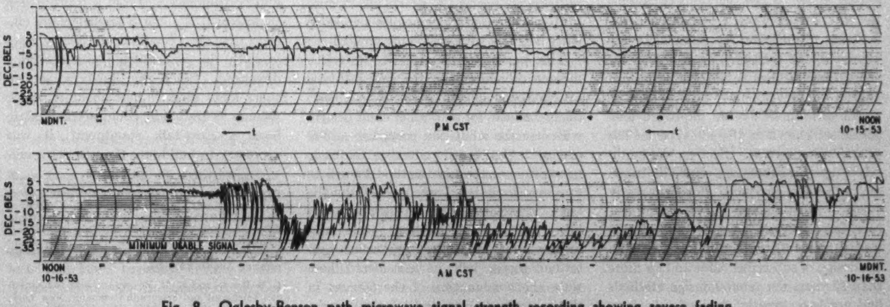


Fig. 8. Oglesby-Benson path microwave signal strength recording showing severe fading

number of cases of automatic switchover to the r-f stand-by where no trouble was found in the main unit. It is felt that part of these was due to intermittent troubles and part to faulty sensing system operation or adjustment. In the multiplex equipment tube failures and faulty wiring have been the main sources of trouble. In the first 6 months of operation, about 35 tubes of the 700 odd in the equipment were replaced due to in service failure or suspected failure (more were removed during installation, precut-over tests, and routine maintenance checks).

Experience with the emergency power equipment has been satisfactory. The few troubles which were encountered have been successfully remedied. It was found that most of the commercial power interruptions were of short duration, several occurring a few minutes apart during storm periods. To avoid having the microwave equipment go through a warm-up cycle with each of these interruptions, the emergency power unit is set to run for about 15 minutes after the last interruption.

The noise on the voice channels of the system is somewhat higher than the objective for channels connected into the Bell System toll network, but is satisfactory in this case. The noise averages 40 db adjusted on the party lines and 35 db adjusted on the 2-point circuits as measured on a *2B* noise measuring set with *F1A* weighting and referred to the zero-db-below-1-milliwatt level. These figures represent a signal-to-noise ratio of about 45 db.

Plans for the Future

At present it appears that the fading effects should be reduced to a point where continuous interruptions of telemetering information over 1 minute in duration occur very infrequently. On this basis, improvement of the Oglesby-Benson path propagation is indicated. It is believed that the path requires a greater fading margin to overcome the effects of the marginal conditions previously mentioned. This is to be accomplished by increasing the antenna system gain at Oglesby by the use of a larger (6-foot diameter) vertically directed parabolic antenna. This is expected to result in a 3-db increase in the fading margin and should reduce the outage time on the path to about one-half of what it would be otherwise. In addition, it is planned to explore fully the use of the differential squelch and associated circuits to overcome fading effects in connection with the operation of the load control equipment.

Conclusion

The operation of this system to date indicates that 6,800-megacycle radio relay can be satisfactorily employed to provide private communications for power companies. Proper design of such systems requires care in selection of paths and antenna heights to avoid such difficulties as strong terrestrial reflections and improper clearances under most meteorological conditions. A fading margin consistent with the desired reliability must be provided. Squelch circuits may be employed to protect telemetering channels during short interruptions. The use of stand-by equipment with automatic switchover, emergency power equipment, and an adequate maintenance scheme are features which can be valuable aids in increasing system availability.

References

1. THE MOTOROLA SYSTEM OF TELEMETER MULTIPLEXING, L. G. **Walker.** *The Complete Papers and Talks of the 1953 National Telemetering Conference*, Chicago, Ill., pp. 213–17.

2. RESULTS OF MICROWAVE PROPAGATION TESTS ON A 40 MILE OVERLAND PATH, **A. L. Durkee.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 36, Feb. 1948, pp. 197–205.

3. SELECTIVE FADING OF MICROWAVES, **A. B. Crawford, W. C. Jakes, Jr.** *Bell System Technical Journal*, New York, N. Y., vol. 31, Jan. 1952, pp. 68–90.

4. PATH TESTING FOR MICROWAVE RADIO ROUTES, **R. D. Campbell.** *Electrical Engineering*, New York, N. Y., vol. 72, July 1953, pp. 571–77.

◆

No Discussion

Simplified Transmission Engineering in Exchange Cable Plant Design

L. B. BOGAN
MEMBER AIEE

K. D. YOUNG
ASSOCIATE MEMBER AIEE

IMPROVEMENT in telephone sets has been responsible for significant advances in the transmission art over the years and has always been an important part of the development program. The recent introduction of an improved telephone set known as the 500-type set has probably had greater effect on the methods of designing exchange cable plant than any other set produced since the early 1920's. Even though improvement had been large in previous sets, transmission was, in general, the controlling factor in the design of subscriber cable and for more than 25 years the general design methods had not been greatly changed. In fact, until very recently, principles in use as early as 1928[1] were still being applied.

This paper will show how the transmission advantages of the new 500-type set have been applied to simplify engineering procedures for subscriber cable design, will describe what the economic stakes are, and give details covering the basic procedures of the new method known as "resistance design," as opposed to the term "transmission design" which is applied to the older method. Only subscriber lines, or as they are called subscriber loops, will be considered here since the combination of the 500 set in the subscriber plant and relatively inexpensive repeaters[2] in the trunk plant have made it practical to design loops and trunks independently.

The 500-Type Set—Transmission Description

The new 500-type telephone set has been described in detail as to construction and performance.[3] From a transmission viewpoint it is about 5 decibels more efficient than its predecessor the 302-type set in both the transmitting and receiving directions. Since the 302-type set on short loops already had about as much volume as the plant could accommodate from a cross-talk standpoint, it was necessary to build into the new set a net-

---

Paper 54-411, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 18, 1954; made available for printing August 23, 1954.

L. B. BOGAN and K. D. YOUNG are with the American Telephone and Telegraph Company, New York, N. Y.

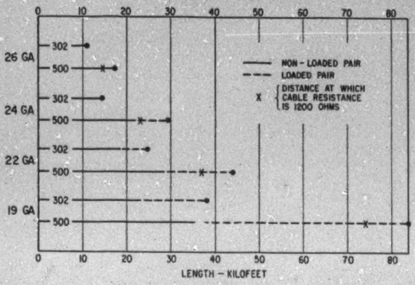


Fig. 1 Maximum lengths of each gauge, loaded and nonloaded, which can be used with 500-type set and 302-type set calculated for same value of transmission loss

work which controlled the increased efficiency from little or none on extremely short lines to the maximum on the longer lines.

For many years, plant design has been based on the older 302-type telephone set and its capabilities are generally well understood. In Fig. 1, for a typical value of loop transmission, the comparable performances of the 500- and 302-type sets are shown in terms of the permissible lengths of each of the wire gauges standardized for use in loop plant. The advantages of the 500-type set become immediately evident by a study of this very simple chart which indicates the following.

1. The 500-type set on any gauge outperforms the 302-type set on the next coarser gauge.
2. With the 500-type set, conductor resistance ranges of central offices, set by supervision and pulsing limitations, control the permissible length of any gauge.
3. Transmission somewhat better than the typical limiting values formerly used with the 302-type set can be provided by the 500-type set even though the conductor resistances were extended to at least 1,500 ohms in the latter case.

This last fact indicated that, to obtain full benefit from the new set, central offices having the lower values of conductor resistance range would require range extensions. Consequently, steps were immediately taken to provide ways and means of extending these ranges, e.g., from around 900 ohms up to about 1,300 ohms in one of the most commonly used types of central office. The possibilities of further extensions in both pulsing and supervision ranges is under study.

Past Design Practices—Meeting the Decibel Limit

In the past, it was the practice in design to assign a transmission value known as a "loop limit" to each operating center. The loop limit was the maximum permissible loss for any loop in the office and was a grade of transmission which it was expected would be obtained only on a small percentage of the loops. The loop limits were determined by economic studies balancing the costs of loops and trunks for certain over-all transmission losses from subscriber to subscriber. These studies involved a process of successive approximations and were rather time-consuming since some exchange areas might include up to 30 or more operating centers. Trial loop limits had to be set for each center and trunk and loop cable plant priced out for these conditions, a new set of limits used and a new cost obtained until, by this cut-and-try process, the most economical combination was obtained.

In addition to setting the loop limit, the cut-and-try procedure was also required in the design of individual loops to stay within the limit. It meant taking into account such items as line loss, loop loading, bridged cable losses, terminal reflections, set efficiencies, and even existing over- or undergauged plant before final choice of cable gauge on any one project could be made. All of these factors had to be evaluated directly in terms of decibels and then juggled so that the total loss stayed within the loop limit. In some cases, the gauge required in the feeder cable would depend to a considerable extent on the length of bridged distribution cable as well as on the over-all length of the loop.

Such a complex situation encouraged the designing of cable plant to meet a particular set of short-lived circumstances rather than a long range over-all plan. With enough experience, it was of course possible to produce good cable designs, and it was not unusual in an area to carefully train one man or a small group of men to do the transmission portion of designs rather than to try to give all the field engineers the knowledge they needed to do the job. This not only created an air of mystery and distrust of the decibel but made for bottlenecks because of the difficulty of arriving at an agreement on the final design. A common measuring stick was hard to find.

Resistance Design—The Simplification of Design Methods

Resistance design made possible by the 500-type set applies a few very simple guides.

1. Select the most economical gauge or combination of gauges permitted by the conductor resistance range of the central office.
2. Use 500-type sets where needed.
3. Apply line loading to all loops which extend beyond a selected distance, usually about 18 kilofeet (kf), using as many points of loading as possible with consideration for correct load spacings.
4. Limit the length of bridged cable in total to about 6 kf.

With resistance design, the engineer deals with ohms which are easier to handle than decibels in view of the number of interrelated transmission items involved. The choice of gauge depends only on the total cable resistance and can be done on a cable route basis. The method provides a common measuring stick for all engineers and rechecking is easily accomplished. Intimate knowledge of the transmission art is not required, yet the procedures, if followed, will assure that the transmission results will be good. Since the design involves only physical quantities of cable, i.e., gauge and length, it can be completed by the field engineer who is most familiar with the field conditions, and the transmission engineer is allowed to spend his time on projects where his particular training may be used to better advantage.

In addition, the gauge of cable placed in the loop plant automatically will be the cheapest for the permissible conductor resistance range and the resulting transmission, using 500-type sets and loop loading as specified, will be better on the average than that obtained under the previous complex procedure using the 302-type set. In fact, instead of the construction being a result of a limit, the limit is the result of the construction. This has given rise to the term "natural loop limit," a maximum value of transmission dependent on the signaling limitations and maximum loop length of the central office.

Evaluation of the New Design Method

While the resistance design method is simple, it had to be worked out. Nor

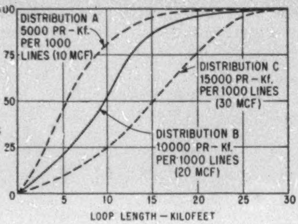


Fig. 2 (left). Typical loop length distributions

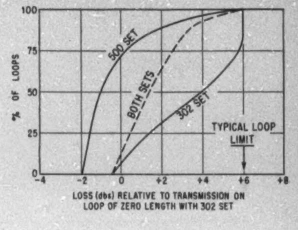


Fig. 3 (right). Loop transmission loss distributions

was it self-evident in the beginning that it would produce greater savings than more complex methods of design giving about the same over-all transmission results. These things had to be evaluated, at least roughly. How this was done, and is still being done in specific studies, is outlined in the following.

With the large number of loops, the variation in cable arrangements and the many office area configurations, the evaluation of the effects of the 500-type set and the new design method required a statistical approach. The study of loop length distributions is of primary importance since they are basic to the transmission provided and to the amount of cable required in an area.

A large amount of data have been collected to determine the nature of loop length distributions as found in the plant. These curves fall into rather general patterns. However, the maximum line length varies considerably since there is no uniformity in size of central office areas, largely because of varying telephone and population concentrations. In some densely populated areas, the maximum length of loop is as low as 10 kf, in others it may be as long as 35 kf. In rural areas, the lengths are even greater. In Fig. 2, for example, the curves show the general shape of the distributions fitted to a maximum length of 30 kf. *A* and *C* define the extremes of the distributions found, and *B* is the weighted average. As an indication of how the character of these distributions affect the amount of cable needed, the million conductor feet of cable required per 1,000 lines, assuming 100-per-cent efficiency, for each case is shown.

Table I. Comparison of Gauge Distribution. Typical 302-Type Set Versus 1,200-Ohm 500-Type Set Design

Distribution	Percentage of Gauge Required With					
	A		B		C	
Set	302	500	302	500	302	500
26 gauge	92	97	59	80	35	47
24 gauge	5	2	21	17	12	29
22 gauge	2	1	17	3	33	24
19 gauge	1		3		20	
Million conductor feet per 1,000 lines	10		20		30	
Tons of copper per 1,000 lines, approximately	5	4	15	10	37	21

Provision of Better Transmission

Using the distribution *B* of Fig. 2 as basic data, the gauge for each loop may be determined and the loss calculated under each design method. The resulting loop loss distributions are shown in Fig. 3. As can be seen, the 500-type set with resistance design will provide equal maximum and decidedly better average transmission than the 302-type set with transmission design. Detailed explanation of gauge determination using resistance design is given in the Appendix.

In the latter case, the practice of trying to keep within a limit has caused the loss distribution to be skewed, increasing the percentage of lines at or near the maximum. As indicated previously, this is not desirable from a transmission standpoint and is avoided under resistance design primarily through the more liberal treatment of loop loading. This design calls for the installation of the maximum practicable number of load points and sets up a definite kf-distance limit for nonloaded loops rather than using a decibel limit. As a result, most of the loaded loops will be somewhat better than limiting and transmission distribution with a better average is produced. In the long run, this loading plan with the 500-type set will not require any more coils than the 302-type set under the old plan, particularly if the requirements of the old plan are considered in the light of the recent extension in conductor resistance ranges.

Continued Use of 302-Type Sets

There are some 30 million 302-type sets already in service and it is obvious that for economic reasons these cannot be discarded especially as they will provide satisfactory transmission service on a large proportion of the loops for many years to come. Also, the transition, in terms of gauge changes, will be a slow process as work is not done every year on every loop and only in those cases where actual gauge changes are made will the 500-type set be a definite requirement. With both sets integrated into the design, it is estimated that over the next 5 to 10 years the distribution of loop loss may approach the dotted curve shown in Fig. 3.

The Treatment of Bridged Cable

The effects of bridged distribution cable has been excluded in these curves. In actual plant, the length of these bridged taps varies greatly with distance from the central office and their effects on transmission are a function of their length, gauge, and location. For general use, it has been customary to assume a fixed loss for each 1,000 feet of tap and the effects of several short lengths are lumped into one tap whose length is the total of all the short ones. This approximation involves only a small error and is adequate for loop design purposes. Studies of bridged tap length per loop show that more than 90 per cent are less than 6 kf in length with the average about 2 kf. Consequently, the 6,000-foot limitation on bridged taps is not expected to be a hardship. At any rate, the distributional effect of bridged taps will be to increase the average loss by approximately 0.5 decibel and the maximum loss by 1.5 decibels.

Significance of Copper Savings

An indication of the relative copper economies permitted by the 500-type set can be obtained by determining the total conductor feet of each gauge of cable required per 1,000 lines, for the distributions shown in Fig. 2, on two bases, "Transmission Design" for a typical loop limit and "Resistance Design" for a 1,200-ohm office. Table I is a tabulation of the

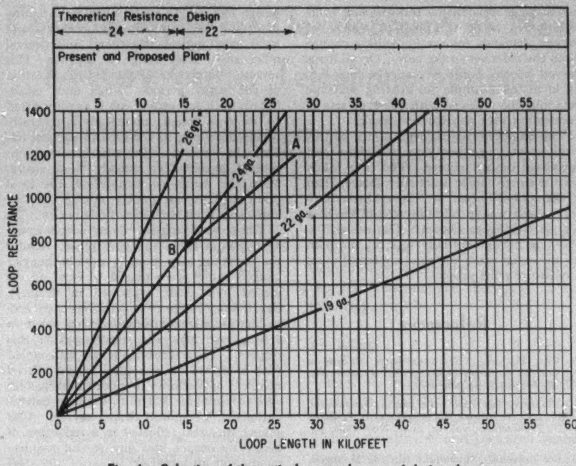


Fig. 4. Selection of theoretical gauges by use of design sheet

results. Note that the advantage accruing from the 500-type set increases as the distribution changes from *A* to *C* of Fig. 2. Within the 30-kf maximum line length shown, no 19 gauge is required with the new set. There are substantial reductions, ranging from 25 to 45 per cent in the tons of copper required, which is a decided economic advantage for the 500-type set. A practical result of the application of the 500-type set has been a decrease of about 0.1 ton of copper per million conductor feet of cable shipped since the 500-type set was first introduced in late 1949. In a typical year, this would be equivalent to some 4,000 tons of copper saved.

Summary

The preceding discussion has shown that, with the 500-type set, conductor resistance range extensions, and the application of resistance design, a threefold advantage is obtained.

1. Engineering is simplified and time is saved.
2. Better transmission is provided.
3. Important economies result.

There are many sides to this subject which are beyond the scope of this paper such as the modification necessary in rural areas having substantial open-wire development, and the zoning plans to insure that the new sets are directed to the loops where they will do the most good. The discussion has touched only the high spots susceptible to generalization and which indicate the fundamentals of the job.

Appendix. Use of Design Chart for Gauge Selection

Basis of Gauge Selection

Under the section entitled "Resistance Design," the length of the loop and the conductor resistance range of the central office determine the cable gauge or gauges selected. A 2-gauge combination such as 26 and 24, 24 and 22, or 22 and 19 will generally be used. Under certain conditions, e.g., conduit congestion on a portion of the route, three gauges might prove more expedient. This can be cared for easily in the design since the length of the third gauge is usually well defined. Having two gauges, a total length and a total permissible resistance, simultaneous equations could be set up and solved to provide answers as to the length of each gauge which could be used. For ease in engineering, a graphical means of solving these equations can be used and may be devised to cover the possible range of gauges and length of loops.

The Design Sheet—An Aid to Engineering

A work sheet has been developed whereby the correct theoretical gauges can be readily determined and on which the engineer can complete his design as well as show all pertinent details. Fig. 4 is a typical sheet. The resistance of each gauge is plotted against length and shown in the heavy lines. To find the theoretical gauge combination for any resistance value and distance, simply draw a line from the intersection of these two values parallel to the adjacent coarser gauge back toward the origin. The intersection of this line with the adjacent finer gauge line determines the length allowed for the finer gauge. The line *AB* in Fig. 4 makes the solution for 1,200 ohms and 28 kf. The theoretical solution can be drawn as a single line diagram in the block provided at the top of the work sheet.

In actual plant, these theoretical solutions are seldom achieved because of a lack of convenient splicing locations, reuse of existing cable gauges, and conduit conditions. The simple single-gauge or 2-gauge combination does provide a basic plan for

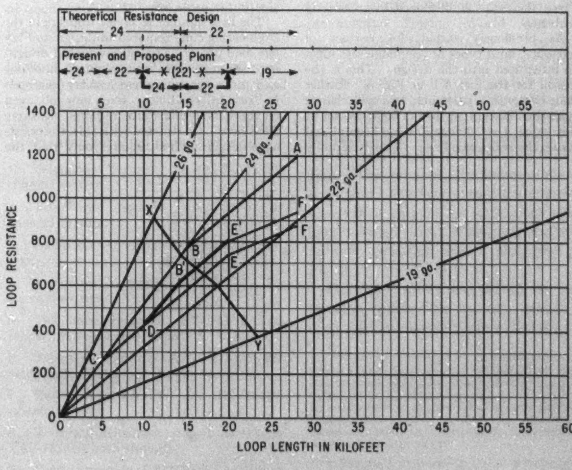


Fig. 5. Graphical solution of typical problem in design

the route, however, and additions or replacement may be easily fitted into the fundamental plan.

Solution of Example Problem for Cable Relief

In this problem the details on which the design is based are as follows:

1. The existing cable run is made up of 5 kf of 24 gauge, 15 kf of 22 gauge, and 8 kf of 19 gauge for a total line length of 28 kf.
2. Cable is to be replaced between 10 and 20 kf.
3. The conductor resistance range of the central office is 1,200 ohms.
4. Loop-loading will be used beyond 18 kf.

As described previously, the theoretical solution for 1,200 ohms and 28 kf is made by drawing in the line *AB* and the theoretical design is indicated in the proper block at the top of the work sheet. The existing cable plan can be laid out gauge by gauge on the body of the work sheet and is the line *OCDEF*. This is done by drawing lines parallel to the proper gauge line at the proper kf-distances starting from the origin.

In the relief section, the gauges to be used are determined from the parallel theoretical design, as shown in Fig. 5. The result is plotted as the line *DB′ E′*. The point *F′* shows the effects of the proposed cable on the point *F*. This can all be summarized in single-line diagram form at the top of the work sheet. Thus, the cable gauges for the replacement cable are chosen so that they will fit into the ultimate design for the entire run. As indicated here, no advantage is taken of overgauged existing cable plant since this may require future expensive rearrangements, or continuation of overgauged plant in cables containing larger numbers of pairs. Even replacement of this cable when future work is done in some other section of the route might be required. Thus, the route is developed in an orderly sequence.

As previously stated, for reasons of economy, both types of telephone sets must be integrated into the design. This is the reason for the line *XY* in Fig. 5. Rather than compute a particular loss and change from the 302-type set to the 500-type set at this value, a set change line has been placed across the resistance lines. This is actually a contour of a particular transmission value determined to be the point at which the service demands require immediate change from the older set to the new. On all loops served by this cable beyond the line *XY*, 16 kf in this example, all existing 302-type sets must be replaced with 500-type sets in connection with construction work.

The engineer can through study of this work sheet, find other ways to make use of its visual design picture. The time taken for the gauge choice is actually shorter than it takes to tell about it. Several routes to serve an area may be explored and the best for all conditions chosen. The design for this one project actually sets up the plan for the entire route and can be referred to on any future relief project.

References

1. The Planning of Telephone Exchange Plants, **W. B. Stephenson.** *AIEE Transactions*, vol. 47, July 1928, pp. 809–18.

2. Negative-Impedance Telephone Repeaters—Application in the Bell System, **A. F. Rose.** *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 430–35.

3. An Improved Telephone Set, **A. H. Inglis, W. L. Tuffnell.** *Electrical Engineering*, vol. 70, no. 9, Sept. 1951, pp. 770–75.

◆

Discussion

J. M. Moss and **P. R. Ferguson** (Pacific Telephone and Telegraph Company, San Francisco, Calif.): Several years of experience with the resistance design method as described by Mr. Bogan and Mr. Young, in a very large operating area with great diversity of telephone distribution plant, has proved the high value of these techniques. Notable are the saving in field engineering time, the saving in outside plant investment, and the improvement in average transmission in the areas affected.

The first significant saving appears in the engineering time required to design and lay out loop plant. The old method of designing to transmission limits required laborious and time-consuming cut-and-try methods to "zero in" the limit. Each new job, even on the same feeder route, required the same treatment. Contrasted with this, the resistance design method requires only that the design loop length (determined from immediate and forecasted line requirements and local geographical situations) and central office supervisory range be known. This sets up a fundamental plan for the duration of the study period. When once established, each succeeding job is readily fitted to the basic design; no restudy is generally required as long as the growth pattern follows the forecast. The solution is accomplished graphically, quickly, and easily. With the design so established, the local plant engineer can quickly pick up the plan and carry it on with very little review.

The second tangible saving occurs in investment in cable and wire plant. Under the limiting loop plan, generally heavier gauged cables were placed to meet an assigned loop limit expressed in decibels. Supervisory range considerations were secondary and usually presented no problem as cables were generally overgauged in this respect. The advent of the 500 set coupled with extended central office ranges offered excellent opportunity to reduce cable gauges and to simplify design to 2-gauge combinations with resultant copper saving. This saving was also reflected in a reduction of heavy messenger and guy strand requirements and an increased usage of 109 steel wire where much copper-steel and even copper wire had formerly been required.

Transmission improvement has resulted due to the application of 500 type sets and full loop loading, particularly in areas from 3 to 5 miles from the central office. These are the areas where most of the large subdivision development is taking place. It is estimated at this time with both 302- and 500-type sets in plant that about 25 per cent of the loops will show transmission improvements of 2 to 3 decibels. Considerable reduction in the number of customers receiving limiting transmission has also resulted.

In conclusion, the following immediate benefits are apparent:

1. Field engineers are enthusiastic over the plan and have so indicated by wholehearted acceptance of its principles.
2. Administrative and investment costs are being reduced and will probably continue to show savings as the plan becomes more firmly established.
3. Transmission improvement has resulted with fewer transmission complaints on exchange loops which require individual study.

Considerations for Development of New Military Carrier Telephone Systems

R. S. BOYKIN
NONMEMBER AIEE

J. H. JOHNSTON
NONMEMBER AIEE

S. D. BEDROSIAN
NONMEMBER AIEE

EXPERIENCE gained by the Signal Corps during World War II and the ensuing experimental period formed the basis for development of the new military 4-channel and 12-channel carrier telephone systems. The heavy, stopgap equipments of World War II while giving reliable service were found to be far from ideal for military use. Experimental work established the considerations which now govern military carrier telephone design.

The trend to high-speed warfare dictated the expansion and increased flexibility of military carrier telephone networks. The 4-channel system of the past has become inadequate for major routes. These routes will be covered by the new 12-channel system. The new 4-channel system will form the forward arms of the carrier network. The result will be faster and more efficient service over greater distances than heretofore possible in military operations. These systems have introduced military carrier telephone service to the type previously obtainable only by more complex commercial systems. This has been done with due regard for economy of manufacture, simplicity of operation, and ease of field maintenance. The result is a pair of highly efficient military carrier telephone systems.

In the summer of 1945, a committee was formed at the Signal Corps Engineering Laboratories to make plans for a completely new military communication system. The report of this committee was completed in October 1945. The report contained the requirements for carrier equipment, spiral-4 cable, switchboards, telephones, and other key items for the new system. Most of these items are now in production or are in advanced stages of development. The requirements for the new equipment to be used in the military long-distance telephone network were based upon experience gained in World War II. The *CF* carrier, radio relay, and spiral-4 cable were used extensively and were most satisfactory. It was definite that more of the same type of equipment was wanted, but smaller and better and in larger quantities. The 4-channel and 12-channel systems under discussion are products of this plan.

The reasons for designing the systems in the way in which they were done and the military necessities which dictated the requirements are discussed in the following.

Logistics

First in importance is probably the matter of logistics. It is, of course, extremely important that this equipment be highly miniaturized and packaged so that it can be handled easily by field personnel. This has been done; almost all items of new equipment are around 100 pounds per package or less, and none ever weighs over 200 pounds. This makes it very easy for one or two men to load or unload the equipment on vehicles and to set it up easily for operation in almost any required location. Fig. 1 shows a comparison of the old carrier telephone terminal *CF-1* with the telephone terminal *AN/TCC-3* of the new 4-channel system.

In addition to size and weight requirements and of even greater importance, from a logistic standpoint, is flexibility. Each item must be able to fill a multitude of needs and meet many varying situations so that an entire long-distance telephone system can be built from a few items of equipment. This flexibility and versatility has been achieved. Ten Signal Corps stock items constitute the entire long-distance telephone network. These items are listed in the following.

1. Telephone terminal *AN/TCC-3*, 4-channel system.
2. Telephone repeater *AN/TCC-5*, 4-channel system.
3. Radio set *AN/GRC-10*, 4-channel system.
4. Telephone terminal *AN/TCC-7*, 12-channel system.
5. Telephone repeater *AN/TCC-8* (attended), 12-channel system.
6. Telephone repeater *AN/TCC-11* (unattended), 12-channel system.
7. Radio set *AN/TRC-24*, 12-channel system.
8. Telegraph-telephone signal converter *TA/-182/U*, both systems.
9. Cable assembly *CX-1065/G*, both systems.
10. Telephone loading coil assembly *CU-260/G*, 4-channel system.

What has been achieved by these 10 items should be emphasized, remembering that military operations demand large stockpiles of millions of individual items at home and abroad. Having such an important communication requirement covered by only 10 items enables concentration of stockpiles and insures that, in time of war, large quantities of all 10 items are available in overseas theaters. This helps greatly to eliminate the risk of having a few missing pieces of equipment put a severe handicap on the entire long-distance communication system.

Ruggedness

In addition to the requirements of miniaturization and packaging for ease of handling, the equipment must be engineered for global use, i.e., it must be rugged enough to withstand shipping and rough handling in the field and still perform in any part of the world. The equipment is, therefore, required to pass a series of severe mechanical and environmental tests to simulate these conditions.

Military Switching Plan

Second to logistics, in order of importance, are the over-all electrical requirements of the equipment itself. The Signal Corps has spent considerable time and contract funds in studying the military telephone system requirements. Much of this work has gone into study of possible transmission and switching improvements. During World War II, with ringdown trunks and 6-, 9-, and 12-decibel (db) transmission losses per trunk, telephone calls over more than two trunks in tandem were made only with the greatest of difficulty. New military switching plans using new equipment call for up to six trunks to be switched in tandem. This requires a far better grade of trunk and a new type of switchboard with supervision capable of setting up such long switched calls. A diagram of a new military switching plan is illustrated in Fig. 2.

This figure shows that a 3-db net loss trunk is required for the new switching plan. To achieve this low net loss of 3 db

---

Paper 54-408, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 18, 1954; made available for printing August 31, 1954.

R. S. Boykin, J. H. Johnston, and S. D. Bedrosian are with the Signal Corps Engineering Laboratories, Fort Monmouth, N. J.

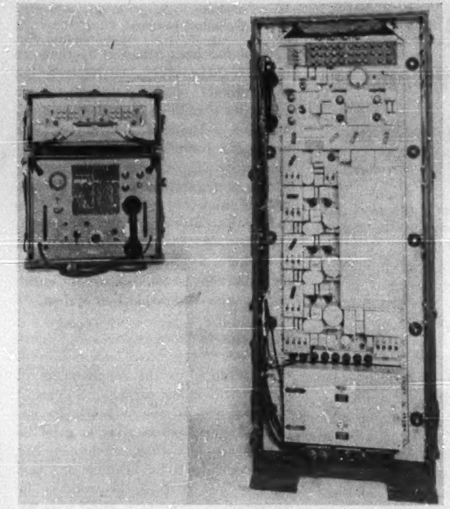


Fig. 1. Telephone terminal AN/TCC-3 compared with telephone terminal CF-1B

requires good regulation and control of the transmission losses and gains in the carrier systems and radio sets. If this is not done, the trunk circuits are in danger of instability and high distortion when the net loss is allowed to decrease to near 0 db. For this reason, the permissible net loss variations in the new equipments are held to ±2 db. Another important characteristic is the bandwidth of the telephone trunk. The new equipment provides a trunk with a bandwidth of 350 to 3,450 cycles (3-db points). Fig. 3 shows the difference between one of these trunks and five of them connected in tandem. To have a good circuit to talk over, after five or six trunks are switched together plus the user's loops at each end, it is essential that each trunk have a wide bandwidth, as shown in Fig. 3.

The allowable noise on the long distance trunk is 38 dba measured at or referred to 0 db transmission level; dba is the adjusted value of noise as used in the Bell System. The interchannel crosstalk loss is required to be 50 db, or greater, measured at equal transmission levels. The near-end crosstalk loss, or crosstalk side-tone loss, is required to be 25 db or greater, measured at equal level points. The differences in carrier supply frequencies used are required to be such that no channel in a carrier system causes more than 2 cycles change in any frequency transmitted over the channel. All of these electrical requirements must be met over a wide variety of temperature and humidity conditions and fluctuations in a-c line voltage.

Size and Length of System

For communication purposes, a theater of operations may be divided into rear area, intermediate area, and forward area. The 4-channel system is intended primarily as forward-area equipment but may be used to provide tributary or terminal trunks anywhere in the theater. Consequently, it is expected, for the most part, to be used as a very-short-haul system within the division with very-high-frequency radio relay sets or up to 25 miles of loaded spiral-4 cable. The maximum length as a cable system with carrier repeaters is 100 miles. The 25-mile repeater spacing is dictated by noise requirements rather than crosstalk in the cable. With a single repeater, however, the system can be 70 miles long. The terminal consists of two transit cases with a total weight of 175 pounds and 5.7 cubic feet of volume. The attended repeater is in a single carrying case with approximately half the size and weight.

The 12-channel system, on the other hand, is primarily for the intermediate areas from division rear and, in fact, will be the main source of the 3-db long-distance trunks. The maximum length as a cable system, with attended and unattended carrier repeaters, is 200 miles. The unattended repeater spacing is $5^1/_2$ to 6 miles which is limited by near-end crosstalk in the nonloaded cable. The nominal attended repeater spacing of 40 miles is based on several considerations, including the matter of power feed to unattended repeaters, the order wire circuit, and the need for mopup on cable equalization. The 12-channel terminal is 4.4 rather than 3 times the size and weight of the 4-channel terminal. This is because of such factors as complexity occasioned by wider bandwidth, the refinement of providing automatic regulation, and the need to feed power over the cable. The attended repeater is 7/10 the size and weight of the terminal, while the unattended repeater is less than 1/10 the size and weight. These data are summarized in Table I with the superseded World War II system added for comparison.

In line with the stress on flexibility, the forward-area carrier installation can be replaced by a 12-channel system without need for complete replacement of outside plant. It can be set up initially as a radio link while the line crew goes down the spiral-4 cable removing the loading coil assemblies every 1/4 mile and inserting unattended repeaters at the $5^3/_4$-mile intervals. Furthermore, there has been enough margin designed into both the 4-

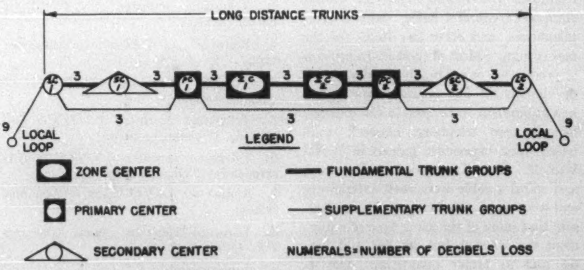


Fig. 2. New military switching plan

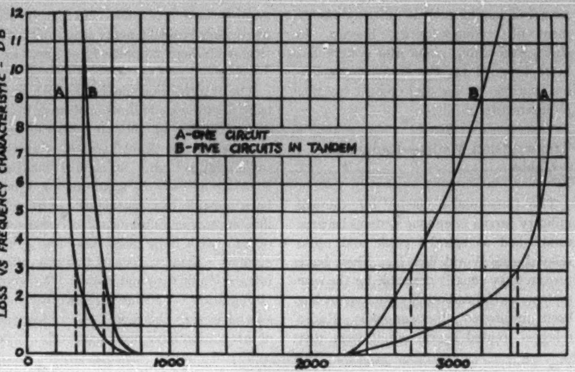


Fig. 3. Effect on frequency band of connecting circuits in tandem

and 12-channel systems to permit emergency operation over captured or reclaimed cable, or open wire lines.

Maintenance

Maintenance of equipment is another prime consideration in the design of military equipment. Every reasonable step must be taken to ease the maintenance problem. Operating systems must be maintained as efficiently as possible and defective components which are too expensive to be thrown away must be repaired. Here again, time must be considered. The whole maintenance program must not require excessively long training periods for personnel; therefore, trouble location and repair procedures must be made very simple and straightforward. Time is also of importance in the field. Trouble must be located in operating systems and be corrected as quickly as possible. Operating crews must be alerted by audible and visible alarms and must be provided with the necessary trouble-locating equipment and an adequate supply of replacement parts. Since time is especially important here, trouble-locating procedures must be streamlined and parts replacement must be easily done. This problem has been met by several methods. The first thing was to provide the most reliable circuits that could be devised. This reduced the frequency of trouble. Then the systems were equipped with a sufficient number of alarms to alert operative personnel of total failure or impending total failure. The next step was the provision of built-in test equipment with features not just for system lineup but for trouble location as well. There is even a push-button tone test to locate defective unattended repeaters and another test to locate certain types of trouble before it appears in traffic circuits. Trouble-location tables were inserted in technical manuals. This completed the features for streamlined location of troubles. Replacement of defective parts has been eased by the provision of plug-in subassemblies. This is a relatively new idea for equipment maintenance as far as the Signal Corps is concerned and, at the moment, there has not been adequate provision in the supply and maintenance organizations to handle this method of maintenance. The plug-in subassemblies make the task of defective parts replacement a very simple matter. Once supply and maintenance are properly oriented, the personnel on system location will not be burdened with the repair of subassemblies. An adequate supply of good subassemblies will be kept on hand and subassemblies which are defective will be returned to rear areas for repair and reissue.

Several points must be settled to realize the full benefit of this plug-in construction. Arrangements must be made between supply and maintenance to assure rotation of subassemblies. Who should handle the rotation? How should the repairs be made? Should running equipments be used as standards or should jigs and fixtures be used? How universal can jigs and fixtures be made? Will there be a few universal items, or will this method have to be abandoned because of an unwieldy assemblage of specialized fittings? Present indications are in favor of jigs and fittings. Supply has another problem. How are these subassemblies to be stocked? How many are to be ordered? How to account for reconditioned items? Eventually these questions will be answered and the full benefits will be realized whereby the operating teams will concentrate upon keeping the systems in operation and the rear areas will repair items for reissue to the operating teams. Thus, the maintenance mission will be fulfilled.

Training for Installation and Operation

It is highly important in times of emergency to get equipment and personnel into the field in the shortest possible time. Equipment may be stockpiled, then quickly transported. Manning the equipment with trained personnel is quite another thing. Personnel must be drawn from all walks of life, trained quickly, and sent forward to fill in the spreading situation. Expansion occurs at a rate far beyond that experienced by a commercial plant. This situation is met by designing equipment and systems that require the least possible training of field personnel. This is done by the study of existing circuits and then devising simple circuits to replace those of the more complex type. When complex circuits must be used, the complexity is "engineered" into the equipment components. For example, a complex network is encased as one complete component and becomes a replaceable item. Of course, this breakdown and packaging of complex-circuit components is based on expected life, cost of component versus cost of re-

Table I. Comparative Equipment Data

System	Equipment Terminal	 Repeater	Weight, Pounds	Volume, Cubic Feet	Relative Size and Weight	Number of Transit Cases
Old 4-channel system	*CF-1*		475	20.3	1.00	1
		CF-3	225	7.7	0.43	1
New 4-channel system	*AN/TCC-3*		178	5.7	0.33	2
		AN/TCC-5	86	2.6	0.16	1
New 12-channel system	*AN/TCC-7*		746	25.2	1.40	9
		AN/TCC-8	511	18.4	1.00	5
		AN/TCC-11	70	1.8	0.12	1

pair, and the resulting reduction of out-of-service time. Thus, training of personnel does not involve complex-circuit analysis. Installation of equipment is based on a few simple rules. Equipment comprised of several units include plug-ended interconnecting cables with all plugs and receptacles clearly marked. No soldering is required and multiple choice wiring does not exist. System line-up adjustment is equally simple. Built-in transmission-measuring equipment provides the means for testing circuits. Line-up controls are clearly marked and are the minimum number required. The majority of line-up adjustments result in a 0-db-reading on the equipment meter. This relieves the personnel of having to remember an assortment of readings during line-up. In short, everything about the system has been designed for the simplest possible installation and line-up to reduce training requirements.

Transmission Media—Radio and Cable

The transmission media for tactical military carrier telephone systems became established as spiral-4 cable and radio relay during World War II. These have proven to be equally desirable for the new systems. The radio relay equipment has been improved tremendously. Radio relay is treated as though it were just another section of cable. In fact, its use has replaced cable in many situations, particularly where mobility is of great importance.

The transmission media are both basically 4 wire which allows for transmission and reception at the same carrier frequencies in the cable system and the radio relay system. By careful attention to design and system requirements, it has been possible to construct the extremely flexible carrier equipments necessary to operate with both radio relay and cable systems. Some years ago, this was called integration of wire and radio. The word "integration" became shopworn from excessive use; however, the objectives of integration have been achieved.

No Discussion

A New Cable Design for Military Carrier Telephone Systems

H. F. X. KINGSLEY
NONMEMBER AIEE

A COMPANION paper[1] has shown how the trend to high-speed warfare has dictated the expansion and increased flexibility of military carrier telephone networks and how the Signal Corps Engineering Laboratories planned and developed a completely new military communication system based upon experience gained in World War II. This paper describes the work done in connection with one important component of the system, namely, a cable for use with 4-channel and 12-channel carrier telephone equipments. As in the design of other military equipments, the cable design goals stressed advantages in weight, bulk, usable bandwidth, and reliability along with significant improvements in electrical and mechanical characteristics in comparison with the carrier cable of World War II. This World War II cable *WC-548* was a rubber insulated and jacketed 4-conductor cable of the star quad type, and it became known as spiral-4.

In April 1946, a contract was placed with the United States Rubber Company for the development of two carrier cables: a lightweight nonloaded cable for use in the proposed 12-channel carrier telephone system; and an ultralightweight loaded cable for use in the proposed 4-channel system. Development work continued into 1950, although basic cable design work was completed in less than 2 years. It might be well to trace some of the general outlines of the work at this time, since this will make unnecessary certain deviations from a direct path later.

Two Cables

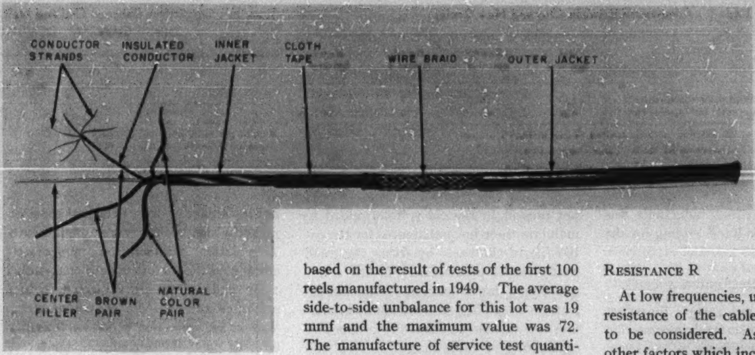
As previously pointed out, the original concept called for the design of two cables. At the time, this appeared to be absolutely necessary for a number of reasons. It was felt that it was possible to meet all the requirements for the 4-channel system with loaded cable. The wide frequency range required for the 12-channel system and the necessity of providing a reasonable repeater spacing made it appear impossible, however, to consider the 4-channel cable for 12-channel use. It might also be mentioned that thoughts at the time envisioned the loading coil incorporated in the 4-channel cable connector which would result in cutoff well below the top frequency required in the 12-channel system.

The aforementioned concept was radically changed with the advent of the unattended repeater for the 12-channel system which set reduced goals on repeater spacing and attenuation. As work progressed on cable designs, it became apparent that the main factor militating against a single cable was the side-to-side capacitance unbalance limit that could be attained. Since no quantity production of cable is accomplished during a development phase, it was necessary to fall back on data relating to old spiral-4 cable and to estimate from that data what could be expected, or hoped for, in the new cable.

Calculations showed that the 4-channel system would require an unbalance no higher than 60 micromicrofarads (mmf) rms per 1/4-mile length and the 12-channel system no higher than 35. Production data on the old spiral-4 cable indicated rms value of no higher than 125. It was certain that the 60-mmf requirement for the 4-channel system could be achieved in production and reasonably certain that a large part of production would conform to the 12-channel requirement. At this point, which was not much more than 1 year after development had started, it was decided to pursue a single ultralightweight cable design which would be loaded for 4-channel use and nonloaded, but using unattended repeaters, for 12-channel use. Selection of low unbalance reels would be made in

Paper **54-409**, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 18, 1954; made available for printing August 31, 1954.

H. F. X. KINGSLEY is with the Signal Corps Engineering Laboratories, Fort Monmouth, N. J.



Courtesy of Signal Corps Publication Agency

Fig. 1. New spiral-4 cable

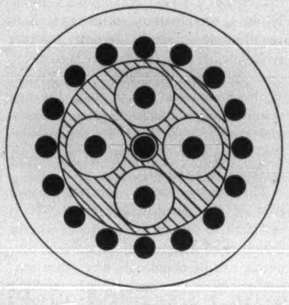


Fig. 2. Cross-sectional view of new spiral-4 cable and details of construction

1. Core, polyethylene, 0.031 inch ± 0.003 inch
2. Strand, 7/0.0136-inch copper, $^9/_{16}$-inch left-hand lay
3. Insulation, polyethylene, to 0.076 inch ± 0.003 inch diameter
4. Cabling, four insulated conductors around core, 2-inch right-hand lay
5. Inner jacket, polyethylene, to 0.233 inch ± 0.005 inch diameter
6. Tape, carbon-filled cloth, 0.0045 inch thick, helically or longitudinally applied
7. Braid, type-302 stainless steel, 16/0.015 inch, 5.75 pics per inch
8. Outer jacket, Buna N-vinyl or vinyl, to 0.370 maximum diameter

production for use in the 12-channel system. Thus, while a single cable was considered feasible, two cable assemblies would be furnished to the field—one for each application.

One Cable

The final step in the process, the decision to use a single cable assembly, was based on the result of tests of the first 100 reels manufactured in 1949. The average side-to-side unbalance for this lot was 19 mmf and the maximum value was 72. The manufacture of service test quantities, 580 reels, was final proof since this lot averaged 8.6 mmf. To use a single cable, it was then necessary to provide loading coils as separate units for use in the 4-channel system.

Cable Design

At this point, it would seem appropriate to revert to the beginning of the development program and cover the specifics of the cable design problem. The two most important elements of a cable are its conductor and its insulation. The conductor, of course, had to be copper. A survey of insulation materials quickly showed that polyethylene, as opposed to any other practical material, resulted in a cable with the lowest attenuation for given specific physical dimensions. Thus, with the conductor and insulation materials decided upon, and with the configuration of a star quad retained from the old spiral-4 cable, the basic cable design took definite shape. As all of the 40 cable constructions which were manufactured and evaluated cannot be discussed in this paper, a description of the final cable design and an analysis of its constructional details and the part they play in providing mechanical and electrical adequacy are given; see Figs. 1 and 2 for cable details.

A great deal of design effort was expended on keeping the carrier frequency attenuation to a practical minimum and, of almost equal importance, on maintaining the attenuation versus frequency characteristic within very close tolerances. This latter design requirement was necessary since the proper operation of the pilot regulation system in the carrier equipment depended upon precisely maintained cable characteristics. Since attenuation is affected by the four basic cable parameters (R, L, G, and C), each was analyzed in turn.

RESISTANCE R

At low frequencies, usually only the d-c resistance of the cable conductors needs to be considered. As frequency rises, other factors which increase circuit resistance must be taken into account. Among these are resistance increases caused by skin effect, proximity effect, and hysteresis and eddy-current losses due to the presence of any other cable components in the electric field It was found that no great reductions were possible in the losses caused by skin and proximity effects, although attempts were made to reduce the former by insulating each of the seven conductor strands with enamel. When the old spiral-4 cable was analyzed, however, it was amazing to see how greatly hysteresis and eddy-current losses contributed to the resistance at 60 kc. The ratio of the a-c resistance at this frequency to the d-c resistance was measured to be 1.70. If only skin and proximity effects were considered, this ratio was calculated to be only 1.23. The additional losses were caused by the aluminum foil stabilizing tape and high carbon steel braid of the old cable. In the new cable, it should be noted that the stabilizing tape is a high-resistance carbon cloth, somewhat less than 10,000 ohms square, and that the braid is stainless steel of low conductivity, about 2.5 per cent, and of very low permeability. The resulting cable has a 60-kc a-c to d-c ratio of only 1.22 and, since this ratio for skin and proximity effects is calculated to be 1.16, it can be seen that a practical minimum has been achieved.

INDUCTANCE L

Cable inductance is primarily dependent on cable dimensions. Attempts were made to increase inductance to effect continuous loading. In this connection, high-permeability tapes were considered but discarded for a number of reasons including the fact that the permeability of these materials changes with working or stress. Magnetic iron oxide was compounded into the insulation with little success. It was concluded that no feasi-

Table I. Comparison Between Old and New Cables

	Old Spiral-4	New* Spiral-4	New† Spiral-4	New‡ Spiral-4
Side-to-side unbalance, mmf per 1/4-mile reel	125	19	7	6
Side-to-ground unbalance, mmf per 1/4-mile reel	650	285	173	100

*First development lot of 25 miles of cable by United States Rubber Company

†Composite average 1952 production by Rome Cable Corporation.

‡Composite average for 1953 production by Rome Cable Corporation.

Table II. Comparison Between Old and New Spiral-4 Cable

	Old Spiral-4	New Spiral-4
Over-all diameter, inches	0.425	0.370
Weight per mile, pounds	545	390
Breaking strength, pounds	600	700

ble method for increasing inductance was available except the lump loading for the 4-channel system.

CAPACITANCE C AND CONDUCTANCE G

As was pointed out earlier, the effect of these parameters on cable attenuation was reduced to a minimum by the selection of polyethylene for the insulation early in development. Additional reductions in these parameters were accomplished by the addition of a polyethylene inner jacket. Since the inner jacket was added to the cable construction primarily to improve capacitance unbalances and the mechanical characteristics of the cable, this feature will be more fully discussed later in this paper.

Side-to-side capacitance unbalance was mentioned in the foregoing, and this would be a good time to discuss capacitance unbalances further to compare the old and new cables in this respect and to detail some of the reasons for the marked superiority of the new design in both side-to-side and side-to-ground unbalances. A comparison between old and new cables at various stages of development and production is furnished in Table I.

The data show the strides made in the reduction of unbalances. It might also be interesting to note that the production data of the Rome Cable Corporation show that 99.7 per cent of their 1953 production was below 19 mmf for side-to-side unbalance and below 316 mmf for side-to-ground unbalance.

Capacitance unbalances can result from electrical variations in the insulating material or from variations in the dimensions and configuration of the quad. The use of polyethylene instead of rubber for the insulation minimized electrical and dimensional variations. It was found to be entirely feasible to limit diameter variations to 1 or 2 mils and, at the same time, achieve excellent centering of the conductor in the insulation. Moreover, the polyethylene conductors were not nearly as distorted or deformed as was rubber by subsequent cabling, braiding, and jacketing operations. The foregoing factors helped to attain a cable with low unbalances. The addition of the inner jacket to the cable construction, a feature not present in the old spiral-4, aided to maintain these low unbalances for the entire life of the cable by fixing the quad configuration permanently in the interstice-filling inner jacket. These design features which helped to attain and maintain low unbalances, and the production engineering of cable manufacturers who analyzed the effects of production equipment and techniques on cable unbalances have resulted in a cable of remarkable electrical properties. In a paper such as this, it is necessary to subordinate certain design problems for lack of time and space. Two such problems, namely, mechanical cable design problems and connector design will therefore be covered in less detail. It should be pointed out, however, that this does not mean that they are of lesser importance.

Mechanical Design

Mechanical design problems are of prime importance to the man in the field using the cable. It was therefore necessary to evaluate the various cable components to determine the best mechanical construction. Many constructions were made with variations in the conductor lay, the cabling lay, and the braid angle. These were evaluated and features were selected for the final design to effect optimum properties in flex, bend, and impact resistance. While this program was essential, the most important new contribution to mechanical properties was the addition of the inner jacket previously discussed. This inner jacket made the core of the cable compact and greatly improved the mechanical characteristics of the cable, especially its impact resistance.

In even a limited discussion of mechanical characteristics, the problem of the cable outer jacket cannot be omitted. Actually this was a major development within a development. A jacket was required which would perform both a positive and a negative function. The positive function was to provide a rugged, abrasion-resistant outer covering for the cable making it usable over a wide temperature range from −67 to +140 degrees Fahrenheit. The negative function was not to damage or distort the heat sensitive polyethylene cable core underneath. This latter consideration immediately eliminated the use of conventional rubber compounds which are subjected to high pressures and temperatures in the continuous vulcanization process. Fortunately, as part of their development contract, the United States Rubber Company did design a jacket compound, a Buna-N-vinyl combination which met both requirements. More recently, other straight vinyl compounds have been developed and are in production. As a result of this mechanical design work, the new spiral-4 cable is smaller, lighter, stronger, and more rugged than its predecessor. Some comparison between the two cables is offered in Table II.

The connector design, too, was a major part of the development effort. It is sufficient to say that resulting connector design is of a universal type; any connector can be coupled to any other connector. It is easy to couple in daylight or darkness, waterproof, and is excellent electrically. Adding to the foregoing performance characteristics the fact that the connector can be assembled in the cable factory without interplant shipment of cable to a subcontractor, it can be seen that the connector design is a worthy companion to the cable.

Conclusion

The Signal Corps Engineering Laboratories are proud of the results achieved by this development effort. In this review, however, cost was never mentioned. Frankly, it is very difficult to compare the new cable assembly to the old in this respect. For one thing, their uses are different, the old cable assembly was exclusively used for a 4-channel system, while the new is a dual-purpose cable used in both 4- and 12-channel systems. The new cable, besides having greater flexibility, has at least three times the message-carrying capacity. In circumstances such as these, it is common practice to compare costs per circuit mile. On this basis, it would be necessary to multiply the cost of the old spiral-4 by a factor of 3 before making a comparison. Secondly, it is necessary to relate the cost of the old cable assembly in the 1945 production to the cost of the new cable assembly in 1954. This "cost of living" factor is usually taken to be 2.4. Thus, a total correction factor of 7.2 might be considered reasonable. If this correction factor is used, it would be found that the cost of the old cable assembly is more than 1,000 percent higher than the new. This is considered embarrassingly too good to use. Therefore, no use of any correction factors is made and it suffices to say that the old cable assembly in 1945 production cost 60 percent more than the new cable assembly costs today.

Reference

1. Considerations for Development of New Military Carrier Telephone Systems, R. S. Boykin, J. H. Johnston, S. D. Bedrosian. *AIEE Transactions*, vol. 73, pt. 1, Nov. 1954, pp. 503-06.

◆

Discussion

E. Mark Wolf (Rome Cable Corporation, Rome, N.Y.): Mr. Kingsley has very ably pointed out the rather dramatic improvement achieved in side-to-side and side-to-ground capacitance unbalance of modern spiral-4 cable. This improvement came about solely as a result of production experience and gradual refinement of manufacturing techniques. No changes in materials, construction, or basic manufacturing methods were involved. Actually, manufacturing costs during 1953 were appreciably lower than during 1952, and individual lengths of cable outside specification limits were virtually unknown.

It is interesting to note how precisely the test data obtained throughout this manufacturing experience coincided with current theories of statistical quality control. The specification limits for side-to-ground capacitance unbalance are as follows:

Maximum average for 100 reel lots: 450 mmf per 1/4-mile length
Maximum individual 1/4-mile reel: 1,250 mmf per 1/4-mile length

The original design limits established as a result of experimental development work in another plant were 280 mmf maximum average for 100 reel lots, and 1,000 mmf maximum for any individual reel. These limits were increased to the present values in the specification as a result of limited production experience, and therefore represented a compromise between what was desired and what could be achieved.

Our early production gave test results seeming to indicate that the specification limits were going to be difficult to meet. Our 100 reel average for all production was running 440 to 455 mmf, with some individual reels measuring as high as 2,000 mmf. At this point frequency distribution plots of individual test results gave a broad rectangular curve bearing little resemblance to the familiar bell-shaped curve to be expected. The shape of this curve gave silent testimony to the fact that the process was out of control. As we searched out and corrected various assignable causes the shape of the curve began to change as the average shifted downward. Toward the latter part of 1953, an analysis of 4,500 consecutively produced lengths showed the following distribution of side-to-ground capacitance unbalance values:

Average of 4,500 lengths: 85 mmf per 1/4 mile
Highest average for 100 consecutive lengths: 104 mmf per 1/4 mile
Number of lengths over 650 mmf: 0
Number of lengths over 450 mmf: 7
Number of lengths over 350 mmf: 32
All other values less than 350 mmf

Needless to say, values such as these plot into a very satisfactory frequency distribution pattern, and reflect the result of experience gained in the course of production.

New Military Carrier Telephone Systems Equipment Features

J. P. HOFFMANN
NONMEMBER AIEE

EVEN though the over-all system design has been carefully planned and the transmission circuitry has been ably executed, a communications system is only as useful and reliable as it is permitted to be by the actual equipment. Equipment usefulness is dependent on ability to reach its operational site undamaged, adaptability to conditions of usage, competence in operation, and susceptibility to rapid reactivation if disabled. With these requirements in mind, two multichannel carrier telephone systems have been developed for military communication in tactical areas. These are a 4-channel system for operation over distances to a maximum of 100 miles, and a 12-channel system for heavier traffic and longer haul operation over distances to a maximum of 200 miles. The over-all plan, the transmission media, and the transmission characteristics of these systems have been covered in companion papers.[1,2] Discussion here will be limited to a description of the actual equipment and its design features as they relate to the basic objectives of durability, transportability, ease of installation, and simplicity of operation, maintenance, and repair under any physical or climatic conditions to be encountered from the arctic to the equator.

General Description

The 4-channel system equipment is comprised of a terminal *AN/TCC-3* and a repeater *AN/TCC-5*. The terminal, illustrated in Fig. 1, is composed of two interconnected units, the modem unit and the amplifier-power supply unit. The modem unit contains the modulating equipment for translation between voice and carrier frequencies. The amplifier-power supply unit contains the amplification, carrier frequency supply, power supply, order wire, alarm, and testing equipment. The repeater shown in Fig. 2 is a single unit containing amplification, power supply, order wire, alram, and testing equipment.

This system was developed to replace the then existing 4-channel system, comprised of the familiar *CF-1* terminal and *CF-3* repeater equipments. The old system was developed early in World War II using standard Bell System components and structural techniques, hastily adapted to military requirements. The new system, designed from its inception to meet military requirements, has many transmission and service functions additional to those of the older system and its equipment weight and volume are approximately one-third those of the older system. The relative volumes and

---

Paper 54-410, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11-15, 1954. Manuscript submitted June 18, 1954; made available for printing August 24, 1954.

J. P. Hoffmann is with the Bell Telephone Laboratories, Murray Hill, N. J.

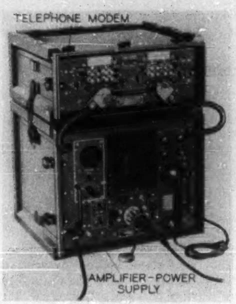

TELEPHONE MODEM
AMPLIFIER-POWER SUPPLY

Fig. 1. Terminal AN/TCC-3

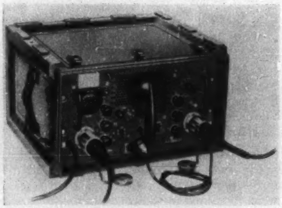


Fig. 2. Repeater AN/TCC-5

weights are listed in Table I.

The 12-channel system equipment is comprised of a terminal *AN/TCC-7*, an attended repeater *AN/TCC-8*, and an unattended repeater *AN/TCC-11*. The terminal, illustrated in Fig. 3, is composed of three modem units, and one each carrier supply unit, subgroup unit, amplifier-pilot regulator unit, order wire-test unit, 200-volt power supply unit, and 600-volt power supply unit, a total of nine interconnected units. The modem units are identical to the modem unit of the 4-channel terminal and perform the same function. The carrier supply unit generates the necessary carrier and pilot frequencies for the system. The subgroup contains equipment for translating the 3-channel modem groups to adjacent frequencies in the band. The amplifier-pilot regulator unit contains equipment for system regulation, amplification, and translation of the carrier band to line frequencies. The order wire-test unit contains the voice order wire, transmission measuring, and trouble location test equipment. The 200-volt power supply unit furnishes direct current for the terminal operation, and the 600-volt power supply unit furnishes direct current over the transmission cable to operate three adjacent unattended repeaters.

The attended repeater, shown in Fig. 4, is composed of one each repeater unit, order wire-test unit, and 200-volt power supply unit, and two 600-volt power supply units, a total of five interconnected units. The repeater unit contains amplification and regulation equipment. The order wire-test unit contains voice order wire, transmission measuring, and trouble location test equipment. The 200-volt power supply unit and 600-volt power supply units are identical to and perform the same functions as those in the terminal.

The unattended repeater, Fig. 5, is a single unit containing the equipment necessary for amplification of the transmission band, together with equipment for cable fault location and by-passing the voice order wire in order to make it independent of transmission band failure. A portable test set *TS-712* and artificial cable unit, normally kept at the attended locations and provided for installation, transmission testing, and trouble location at the unattended repeater, are illustrated in Figs. 6 and 7.

Table I. Comparison of Old and New Systems

Equipment	Volume, Cubic Feet	Weight, Pounds
CF-1 terminal	19.89	475
TCC-3 terminal	5.65	176
CF-3 repeater	7.3	225
TCC-5 repeater	2.63	86

This 12-channel system was designed to supplement the 4-channel system for relatively long haul, heavy traffic routes. Since it is the first of its kind, no direct comparison to an existing system as to volume and weight can be made as was done for the 4-channel system.

Equipment Features

Having stated the basic requirements for and given a brief description of the equipment which makes up these carrier systems, the design features which give them the durability, transportability, flexibility, and operability required for military usage will now be examined.

Terminals and Attended Repeaters

The most obvious feature is the subdivision of the equipment into small interconnected units. The units are entirely self-contained, requiring no separately packed miscellaneous parts for installation. These units range in weight from 63 to 118 pounds and have volumes of 2.0, 2.6, or 3.7 cubic feet. Their small size and light weight achieved by use of miniaturized components, plug-in and bracket assemblies which are easily removable, and aluminum structural parts make them transportable by any means over any terrain and simplify handling by field personnel. Their division along circuit functional lines allows multiple use

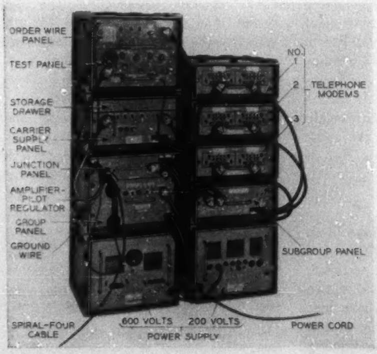

ORDER WIRE PANEL
TEST PANEL
STORAGE DRAWER
CARRIER SUPPLY PANEL
JUNCTION PANEL
AMPLIFIER-PILOT REGULATOR
GROUP PANEL
GROUND WIRE
SPIRAL-FOUR CABLE
NO. 1
2
3
TELEPHONE MODEMS
SUBGROUP PANEL
600 VOLTS
200 VOLTS
POWER SUPPLY
POWER CORD

Fig. 3 (left). Terminal AN/-TCC-7

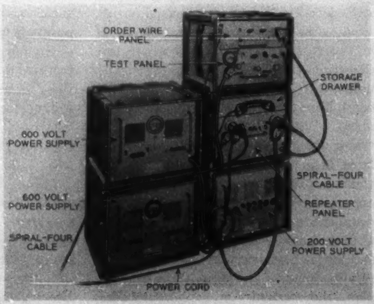

ORDER WIRE PANEL
TEST PANEL
STORAGE DRAWER
600 VOLT POWER SUPPLY
600 VOLT POWER SUPPLY
SPIRAL-FOUR CABLE
SPIRAL-FOUR CABLE
REPEATER PANEL
200 VOLT POWER SUPPLY
POWER CORD

Fig. 4 (right). Attended repeater AN/-TCC-8

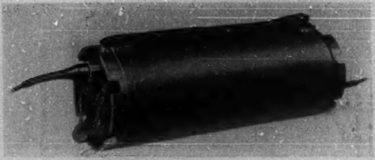


Fig. 5. Unattended repeater AN/TCC-11

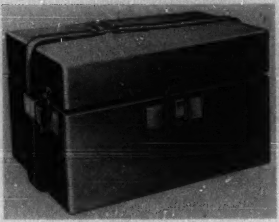
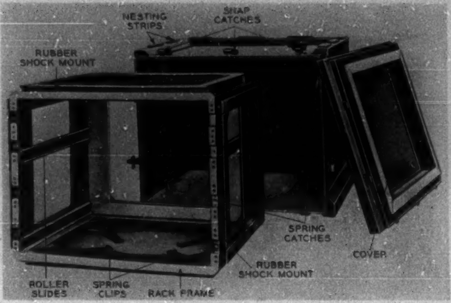


Fig. 6. Portable test set TS-712

Fig. 8 (right). Transit case and shock - mounted framework



in the same or different systems. Manufacture is facilitated and parts reduced by concentrating production on fewer items. Stock-piling is simplified for the same reasons. Subdivision into small units permits control of the amount of heat dissipated within a single unit and removes the necessity of forced-air ventilation in any but power supply units of the 12-channel system. The required ruggedness is built into the components, packaging, and structure.

Each unit, with the exception of the unattended repeater and associated portable test set (which will be covered later), consists of a transit case enclosing components assembled on a chassis in a shock-mounted framework. A typical transit case and framework are illustrated in Fig. 8. The cases are made in three sizes, all 18 inches wide by 20 inches deep but with heights of 9, 12, or 17 inches as required. They are watertight aluminum shells with reinforced edges and gasketed front covers. Interlocking nesting strips along top and bottom edges prevent slippage when cases are stacked. Straps on the sides of the case act as convenient carrying handles during transit and as vertical ties between units in a stack. The framework, with rail-type rubber shock mounts along external edges on all sides except the front, fits into the case and is held therein by spring catches. Tracks on the inside of the frame are located to mate with roller slides on the sides of the equipment chassis. Rail-type shock mounts in the cover of the case bear against the front panel of the equipment chassis and complete the shock isolation of the unit on all sides during transit. A plate in the bottom of the frame equipped with spring clips provides for circuit label storage.

Components, the building blocks for all circuitry and equipment, were carefully selected, not only by the usual criteria of circuit and mechanical reliability but also by size, weight, and shape. They were selected in miniature size where commercially available from military standard approved lists. Where not commercially available to meet the particular requirements of these systems, they were specifically designed to meet the same rigid military requirements in the smallest possible size and most convenient shape. An example is the development, without size increase, of an improved right-angle connector with a larger, more accessible wiring chamber to accommodate the multiconductor cables made up of shielded single and paired conductors. Techniques of plastic encasement and plastic sealing of aluminum encasement were used wherever advantages in size or weight could be realized.

Chassis on which components are assembled are drawer type equipped with roller slides for insertion into the shock-mounted framework. Automatic catches on the slides permit the chassis to be withdrawn only two-thirds of its depth by direct pull. At this point, the catches become accessible and can be manually released to permit complete chassis removal. The automatic catch guards against personnel injury and breakage of connecting cables. Fig. 9 illustrates the feature of partial chassis withdrawal without disconnecting the cables, which relieves front panel congestion by permitting location inside the chassis of line-up adjustments and in-service test jacks. Complete removal of the chassis gives accessibility for maintenance. Although all chassis are basically the same, the detail of design is varied to fit each to its par-

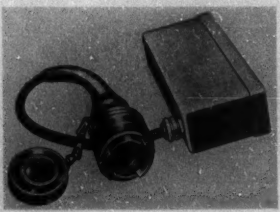


Fig. 7 (left). Artificial cable

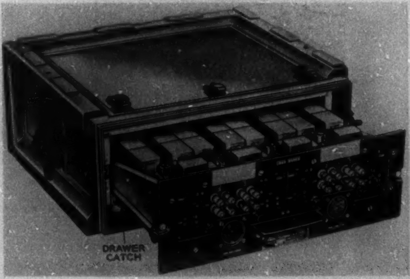


Fig. 9 (right). Typical unit, chassis partially withdrawn

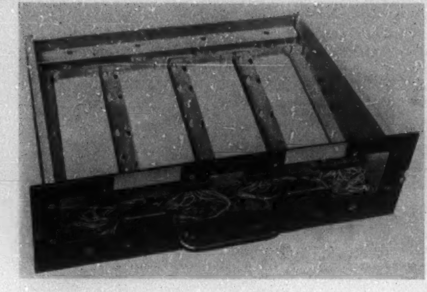


Fig. 10 (above). Channel modem chassis

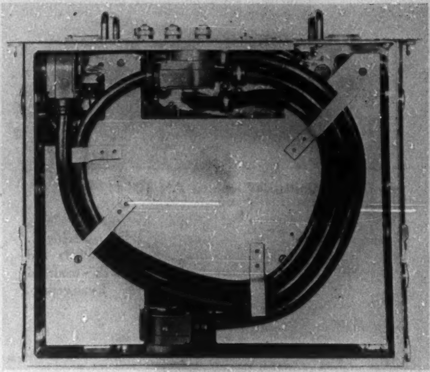


Fig. 12. Cable storage underneath chassis

ticular job. Four chassis have been selected for typical illustration. The channel modem chassis shown in Fig. 10 is of strut and beam construction. This structure gives great strength and rigidity combined with light weight and also provides location devices for the relatively heavy channel plug-in assemblies. Fig. 11 illustrates a conventional dishpan chassis most suitable for single layer component assembly, with ready access to both apparatus and wiring sides. By increasing the depth of the dishpan, a convenient space is made available for the storage of unit interconnecting cable, as illustrated in Fig. 12. A compartmented chassis providing a central ventilated section for power equipment with shielded compartments on either side for transmission equipment is illustrated in Fig. 13. A specialized adaptation of a conventional dishpan chassis to a compact, folded assembly of components packed solidly on both sides of the chassis is illustrated in Fig. 14. Reinforcing channels across the open face of the dishpan serve the dual purpose of strengthening and stiffening the chassis for double loading and also provide support for auxiliary shield plates and plug-in component assemblies. A typical application of the plug-in principle is illustrated here. Removal of plug-in assemblies and shield plates gives ready maintenance access to chassis wiring.

The plentiful use of plug-in assemblies realizes valuable advantages in equipment compactness. Space normally required for maintenance access is utilized for mounting components without in any way impairing accessibility. The property of rapid removal and substitution with a minimum of service interruption eliminates the necessity of on-the-spot repair. Multiple use of the same plug-in assembly in different parts of the equipment aids production and stock-piling and facilitates field reactivation of disabled equipment by reclamation of interchangeable assemblies from other disabled equipment. In these equipments, 15 different plug-in assemblies are used in 51 different places which averages better than three uses for each assembly. It might also be pointed out that the main units and chassis assemblies themselves have these same plug-in characteristics and advantages. Figs. 15 and 16 illustrate the wiring and apparatus sides respectively of typical multiple use plug-in assemblies. A channel modem unit, together with three others in place on their chassis, is illustrated in Fig. 17. This is the largest of the plug-in assemblies and includes its own section of the front panel. Though there are slight variations in components used, the structure of all 4-chan-

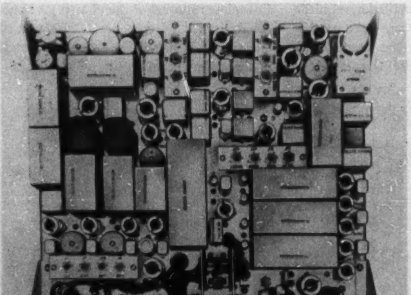


Fig. 11. Typical chassis, dishpan type

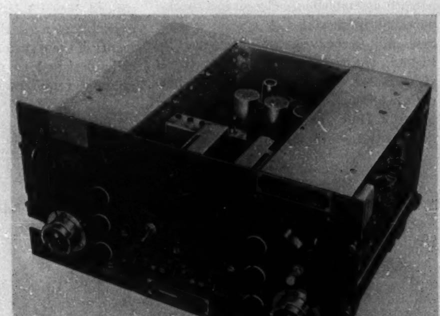


Fig. 13. Typical chassis, compartment type

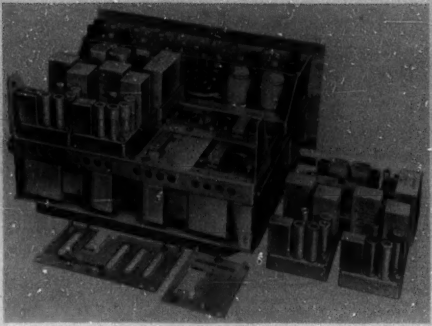


Fig. 14. Typical chassis, folded assembly type

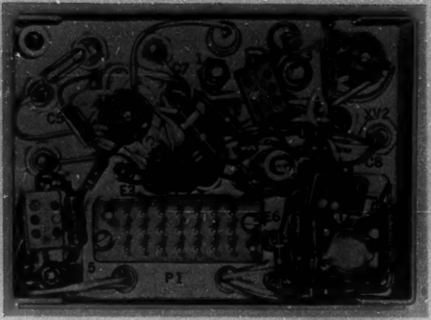


Fig. 15. Plug-in assembly, wiring side

nel assemblies is duplicated for production reasons.

Installation of the equipment is simple and rapid. It requires only the removal of covers, stacking of cases, and plug-in connections of line, power, and unit interconnection cables. Connections between units are made by cables with one end permanently connected and the other end terminated in a plug connector. In use, the cables pass through slots in the containing unit front panel and their connectors mate with panel-mounted receptacles on other units.

UNATTENDED REPEATER

The unattended repeater is designed expressly for its exposed environment where it may lie on the ground or be strapped to a pole without shelter. Its installation is simple, requiring only the plug-in connection of the transmission cables. In case of trouble, another unit is substituted and it is returned to an attended location or depot for repair. It is a cylindrical unit $28^1/_2$ inches long and 10 inches in diameter having a volume of 1.3 cubic feet and a weight of 83 pounds. Its slim, cylindrical shape presents a curved surface less susceptible to injury from missiles or other moving bodies and gives it an unobtrusive silhouette. The casing is an aluminum casting, vacuum-impregnated to make it waterproof, with removable end covers. Removable covers over ports on each end permit access for cable connection and line transmission testing. Webbing straps attached to the outside of the casing act as convenient carrying handles and as a means of pole mounting. The unit with its components disassembled and removed from the case is illustrated in Fig. 18.

The shock mounting of this unit is somewhat unique both as to arrangement and material. Fiberglas cushion material was selected as most suitable to meet the requirements of shock isolation for this unit. This material is composed of glass fibers coated with a phenolic binder and felted. It has little resistance to shear and provides only unidirectional cushioning. Its favorable characteristics are that it has no lateral expansion under compression, is exceptionally resilient, and will return to its original height even after being highly compressed. It has greater vibration damping effect than rubber and returns less energy to the equipment on rebound after impact. It is unaffected by wide temperature variations, its characteristics remaining practically constant over a range of −65 to 165 degrees Fahrenheit. The shock isolation of the repeater was designed to use the material only in compression, thereby taking advantage of its favorable characteristics. Four rails, each supported by three shock isolation cushions in 2-piece overlapping cup mountings buffered to limit shear, are mounted lengthwise in the casing spaced radially 90 degrees apart.

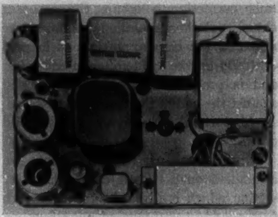


Fig. 16. Plug-in assembly, apparatus side

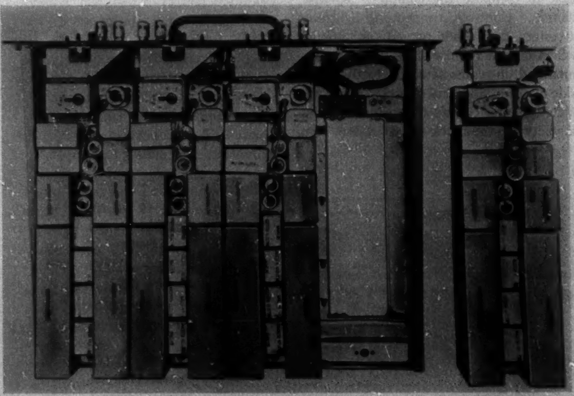


Fig. 17. Channel modem plug-in assembly together with three assemblies plugged into chassis

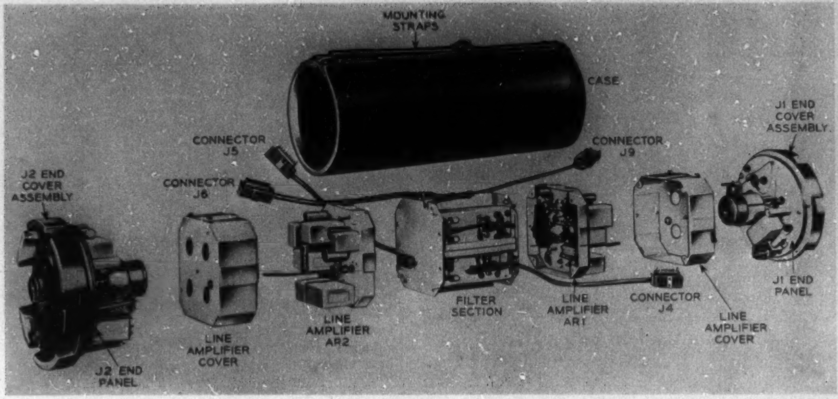


Fig. 18. Unattended repeater, disassembled and removed from casing

The rails are fastened at each end to the wall of the casing and preloaded to obtain the most favorable characteristics of the shock-isolating material and to space the rails correctly to admit the repeater chassis. A shock-isolation cushion in the shape of the circular column is located centrally on the inside of each end cover and is preloaded by being compressed against the end of the repeater chassis assembly when end covers are screwed down.

A number of passive components are assembled on a chassis called the filter section which is divided lengthwise into four quadrants. Identical amplifiers, on dishpan chassis of the same hexagonal outline as the end view of the filter section, plug into both ends of the filter section to compose the complete repeater chassis. Controls, connectors, lightning arresters, and other rugged components which do not require shock isolation are mounted on plates attached to the inner side of the casting end covers. The complete repeater chassis slides into the casing between the shock mount rails. Two flexible cables from each end of the filter section are terminated in connectors, which fit into and are positioned by sockets on the side of the casing to mate with connectors on the end covers. With the end covers assembled, the repeater floats between rails and end cushions and is isolated from shock in all directions.

The portable test set is $13^{3}/_{8}$ inches long, $9^{7}/_{8}$ inches wide, $7^{1}/_{2}$ inches deep, and weighs 19 pounds. It is cased in a waterproof aluminum box with carrying strap and removable cover. The face panel and hinged subpanel compose an integral chassis, illustrated in Figs. 19 and 20. This chassis is set into the case, with the face panel and the control shafts and cable passing through it, gasketed to allow use of the set in driving rain with the cover removed. In the interest of small over-all size, a rugged structural design and the use of a ruggedized meter have permitted the omission of shock isolation in this set.

PERFORMANCE TESTS

To insure that the equipment would perform both mechanically and electrically under global conditions of temperature and humidity after encountering the vicissitudes of wartime transportation, it was subjected to and successfully passed a series of gruelling tests. These tests included a cycling vibration test

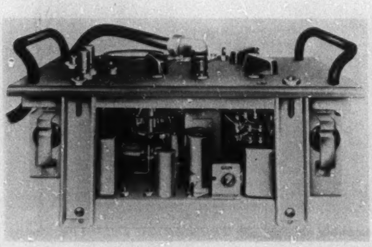


Fig. 19 (left) Portable test set chassis

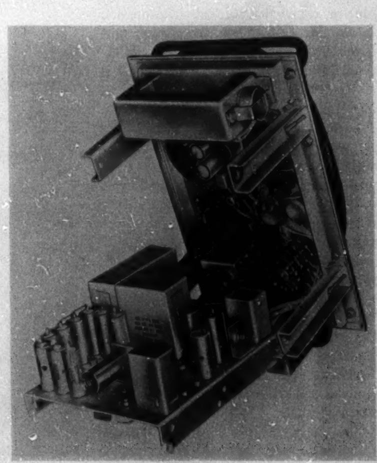


Fig. 20 (right). Portable test set chassis, subpanel hinged open

over a frequency range of 10 to 55 cycles per second per minute with a total excursion of 0.016 inch, a shock test in which the equipment was subjected to shock by forces of 400, 1,200, and 2,000 foot-pounds and a bounce test in which the equipment was subjected to free bouncing for a 3-hour period on a platform mounted on two synchronized eccentric axles moving so that any point on the table moves vertically and longitudinally in a 1-inch circle at a rate of 285 times per minute. In addition, it was subjected to storage temperature extremes of −80 to +160 degrees Fahrenheit and required to perform electrically at full efficiency within the range of −40 to +131 degrees Fahrenheit at 95-per-cent relative humidity.

Conclusion

Careful attention to packaging, structure, weight reduction, operational techniques, and end usage, co-ordinated with intensive circuit development, has resulted in rugged compact portable units, capable of easy installation, operation, and repair. The actual realization of these equipments which are particularly suitable for use in tactical areas makes available the basic building blocks needed for a military communications system.

References

1. Considerations for Development of New Military Carrier Telephone Systems, R. S. Boykin, J. H. Johnson, S. D. Bedrosian. *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 503–06.

2. New Military Carrier Telephone Systems, G. H. Huber, W. F. Miller, C. W. Schramm. *Ibid*, pp. 515–25.

No Discussion

New Military Carrier Telephone Systems

G. H. HUBER
MEMBER AIEE

W. F. MILLER
NONMEMBER AIEE

C. W. SCHRAMM
MEMBER AIEE

Synopsis: Military communication systems have recently been developed for operation in tactical areas using carrier telephone channels on spiral-4 cable. This is part of an extensive program intended to increase the flexibility of the military communications network. Good transmission performance under all environmental conditions has been stressed. These projects make use of a new type of spiral-4 cable. The cable is to be laid in 1/4-mile lengths with or without loading coils at the junctions. A 4-channel telephone system using the carrier frequency range from 4 to 20 kc has been designed for operation on the loaded cable for distances of 100 miles with repeaters at about 25-mile intervals. A 12-channel telephone system using the carrier frequency range from 12 to 68 kc has also been designed for operation on the nonloaded cable for distances of 200 miles with main repeaters at about 40-mile intervals and with "power-over-the-cable" repeaters at about 6-mile intervals. These new low-loss channels make possible the use of multilink connections and the ultimate establishment of a comprehensive military communication network. The circuits have been developed for great reliability, with simple operating procedures, automatic alarms, good maintenance practices, and great flexibility as to usage.

CARRIER telephone technique or multiplex techniques, both frequency-division and time-division applied to wire and radio, are not new to military communication. One of the communication standards of World War II, the *CF-1* carrier telephone terminal, was used with spiral-4 cable, open-wire, and radio in both the European and Asiatic theaters with much success. It was in continued demand so that, when the Korean campaign was in progress, these carrier telephone terminals were again put in production in substantial quantities and were produced until the middle of 1953, when the production of the *AN/TCC-3*, the new version of this carrier telephone technique application, was in quantity production. The *AN/TCC-3* will give better transmission in a much smaller package, with components designed in accordance with the latest military specifications, and is built to provide four telephone-message channels plus one voice-frequency maintenance channel. It is a self-contained, a-c operated, manually regulated carrier telephone terminal. The *AN/TCC-5* is a repeater designed to extend the operation of these terminals up to about 100 miles.

While the 4-channel system was being planned, a thorough study of the military communication needs was in progress. It was determined that, in addition to the 4-channel system, an automatically regulated longer haul system with 12 channels was also needed.

Fig. 1 illustrates a typical section of a probable military communication network with a main route of 12 channels built up of a combination of wire and radio links. The terminals may be simultaneously connected to telephone, superposed telegraph, low-speed picture, or facsimile or other special services, while alternate use of several channels allow high-speed facsimile or data-transmission systems to use the wide band of this main route system. *AN/TCC-7*, the new 12-channel terminal, and *AN/TCC-8* and *AN/TCC-11* repeaters form this new spiral-4 cable carrier telephone system. (*AN/TRC-24*, a new radio set designed for multichannel operation, was also developed as part of this communication plan and is described in a companion paper.[1]) Each terminal and repeater of this 12-channel system is self-contained, a-c operated, and automatically regulated with sufficient precision to permit stable, low-loss operation of all channels with any type of service with up to 200 miles of cable in both arctic and tropic conditions.

The 4-Channel System

The 4-channel system has been designed for operation on the new loaded spiral-4 cable or suitable radio relay links. However, sufficient equalization range and flexibility have been incorporated for operation on open-wire lines and other media.

The spiral-4 cable is loaded by including a 6-millihenry inductance as part of the coupling which joins the 1/4-mile cable lengths. The choice of loading inductance has been made to secure the desired attenuation and also to provide a line characteristic which was amenable to simple line equalization. It was also desired that the shape of the characteristic

---

Paper 54-386, recommended by the AIEE Wire Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 18, 1954; made available for printing August 17, 1954.

G. H. Huber, W. F. Miller, and C. W. Schramm are with the Bell Telephone Laboratories, Inc., Murray Hill, N. J.

As in many projects of this scope, the development has been the result of the efforts of many individuals. It is not possible to give individual recognition to each of these participants. However, the authors wish to acknowledge their contributions, recognizing that they were a necessary and essential part of the successful completion of this project.

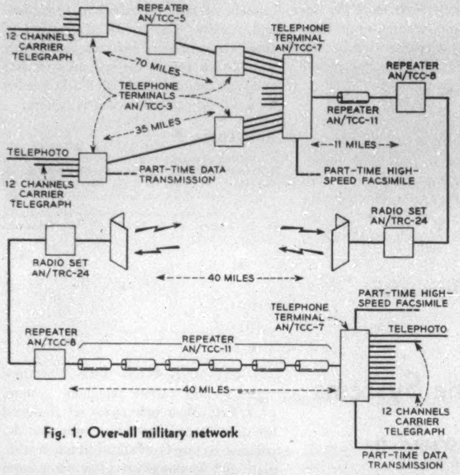

REPEATER AN/TCC-5
12 CHANNELS CARRIER TELEGRAPH
TELEPHONE TERMINAL AN/TCC-7
70 MILES
TELEPHONE TERMINALS AN/TCC-3
REPEATER AN/TCC-8
REPEATER AN/TCC-11
35 MILES
11 MILES
TELEPHOTO
PART-TIME DATA TRANSMISSION
PART-TIME HIGH-SPEED FACSIMILE
12 CHANNELS CARRIER TELEGRAPH
RADIO SET AN/TRC-24
RADIO SET AN/TRC-24
40 MILES
PART-TIME HIGH-SPEED FACSIMILE
TELEPHONE TERMINAL AN/TCC-7
TELEPHOTO
REPEATER AN/TCC-8
REPEATER AN/TCC-11
40 MILES
12 CHANNELS CARRIER TELEGRAPH
PART-TIME DATA TRANSMISSION

Fig. 1. Over-all military network

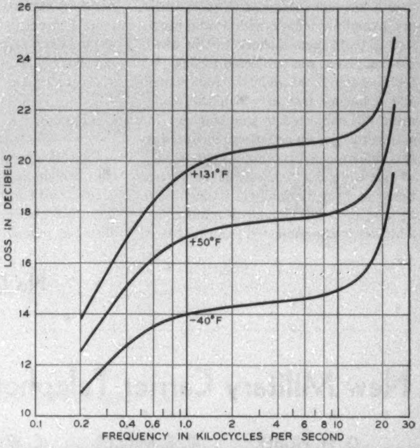


Fig. 2. Spiral-4 cable loss (CX-1065/G). 25 miles of cable with 6-millihenry loading at 1/4-mile intervals

be relatively stable with temperature. These features are shown in Figs. 2 and 3. This gives a transmission medium with a cutoff of about 26 kc and with a loss of about 0.8 decibel (db) per mile at 20 kc.

Four channels using carrier suppressed lower side-band transmission on carriers of 20, 16, 12, and 8 kc, plus a voice-frequency order circuit, are operated in this frequency range. The allocation is shown in Fig. 4. Means are also provided for the use of the entire spectrum for wide-band transmission.

The system may be operated on the spiral-4 cable for distances of about 100 miles with repeaters at 25-mile intervals with normal transmission levels. When conditions warrant, the span can be increased to 35 miles by the operation of a key giving 10 db higher output level.

The terminal (*AN/TCC-3*) consists of two units, and the repeater (*AN/TCC-5*) is housed in one unit. These are shown in Figs. 5 and 6 respectively. The channel modem (modulator-demodulator) unit of the terminal is a basic modulating step applicable to this and other multichannel carrier telephone systems. It also forms part of the *AN/TCC-7* 12-channel system. The second unit in the terminal contains the common amplifying and equalizing equipment, as well as the power supply and auxiliary circuits. Many of these circuits are used also in the repeater.

The block schematic of the terminal is shown in Figs. 7 and 8 and indicates the principal transmission paths. The terminal carrier supply is shown in Fig. 8. It will be seen that the four channels are combined at the output of their transmitting band filters as is customary in frequency-division systems. The combined signals are amplified and sent on to the line. At the receiving end, the line equalizers supply losses complementary to that of the cable, giving a flat input to the receiving amplifier. The channel band filters select the proper channels, the signals are demodulated, and voice amplification gives the desired circuit loss. These channels may be operated 2-wire by the optional use of the 4-wire terminating networks which are included within the equipment.

A number of auxiliary circuits are also necessary to form a complete terminal suitable for operation as part of an isolated communication facility. These are as follows:

1. The order circuit which consists of filtering, limiting, and terminating equipment, a telephone set, a ringer-oscillator for signalling on the order wire at 1,600 cycles, and a series of keys to permit talking and monitoring on any of the carrier channels.
2. The carrier supply which consists of a 16-kc crystal oscillator, frequency-dividing circuits, 4-kc harmonic generator, and output filtering circuits.
3. The power supply containing a filament, plate, and bias supplies derived from 110- or 220-volt a-c input with a tolerance of ±10 per cent. The supplies for both terminal and repeater were designed as nonregulated supplies to save weight and size.
4. Measuring circuits consisting of a test oscillator with detectors at terminals and repeaters and means for applying and measuring test signals at various line-up points in the system.
5. The system alarm consisting of a 4-kc pilot transmitted over the line and tuned detector at the receiving terminal. An alarm is given when the received pilot drops below normal.

Fig. 3 (right). Spiral-4 cable loss (CX-1065/G). 25 miles of cable with 6-millihenry loading at 1/4-mile intervals. Loss change with temperature (relative to loss at 50 F)

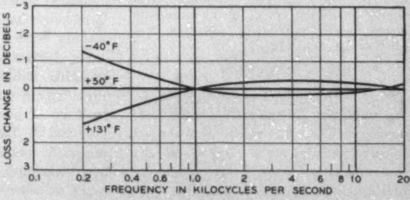
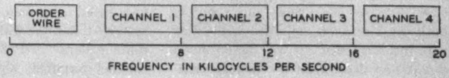


Fig. 4 (right). Frequency allocation of 4-channel system


ORDER WIRE
CHANNEL 1
CHANNEL 2
CHANNEL 3
CHANNEL 4
0
8
12
16
20
FREQUENCY IN KILOCYCLES PER SECOND

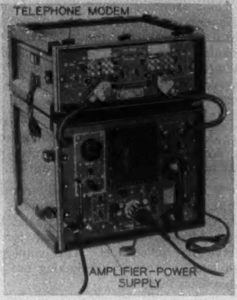

TELEPHONE MODEM
AMPLIFIER-POWER SUPPLY

Fig. 5. AN/TCC-3 terminal: 4-channel system

The block schematic shown in Fig. 9 gives the transmission paths and some of the auxiliary circuits used at the repeater. The circuit treatment is similar to that of the terminal; in fact, many plug-in equipment units are used in common.

Each of the carrier channels provides a high-grade message circuit with a frequency band between 300 and 3,500 cycles per second. Fig. 10 shows a typical band attenuation characteristic and Fig. 11 shows the corresponding delay characteristic. The channel performance is comparable with that of the long-haul commercial telephone channels in use today. It should also be noted that the objective of operating these military channels with a nominal 3-db circuit loss (2-wire) imposes severe transmission requirements on the band shape, the channel stability, and the terminal balances. The objectives are similar to the modern commercial telephone low-net circuit loss operation.

Signalling over the carrier channels is provided by auxiliary Signal Corps equipment. The order circuit also has good transmission performance and a high degree of flexibility so that the attendants can maintain the system with or without radio links and can ring each other even when the circuit losses become quite high. The "special service" transmission is flat to ±5 db from about 4 kc to 20 kc.

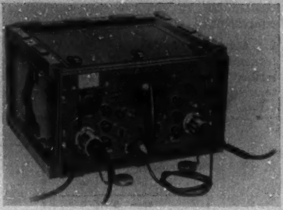


Fig. 6. AN/TCC-5 repeater: 4-channel system

The ability to maintain the transmission of the carrier channels on an "in-service" basis without automatic regulation has been achieved by the system of line equalization involving "flat gain" control at 1,000 cycles. It is thus possible by measuring and adjusting the transmission on the order circuit to hold the carrier channels to their proper net losses.

Other features of this system are:

1. A ruggedized basic modem utilizing plug-in channel units with performance

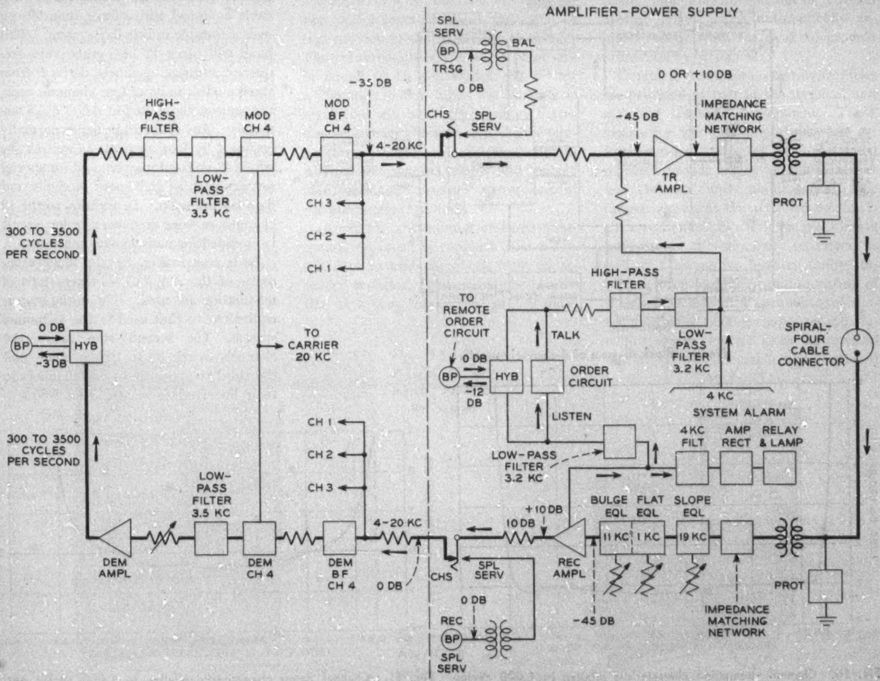

AMPLIFIER-POWER SUPPLY
SPL SERV
BP
BAL
TRSG
0 DB
-35 DB
HIGH-PASS FILTER
MOD CH 4
MOD BF CH 4
CHS
SPL SERV
4-20 KC
0 OR +10 DB
-45 DB
IMPEDANCE MATCHING NETWORK
TR AMPL
LOW-PASS FILTER 3.5 KC
CH 3
CH 2
CH 1
PROT
300 TO 3500 CYCLES PER SECOND
HIGH-PASS FILTER
TO REMOTE ORDER CIRCUIT
TALK
LOW-PASS FILTER 3.2 KC
SPIRAL-FOUR CABLE CONNECTOR
0 DB
BP
HYB
-3 DB
TO CARRIER 20 KC
0 DB
BP
-12 DB
HYB
ORDER CIRCUIT
4 KC SYSTEM ALARM
LISTEN
4KC FILT
AMP RECT
RELAY & LAMP
CH 1
CH 2
CH 3
300 TO 3500 CYCLES PER SECOND
LOW-PASS FILTER 3.2 KC
LOW-PASS FILTER 3.5 KC
4-20 KC
+10 DB
10 DB
BULGE EQL
FLAT EQL
SLOPE EQL
11 KC
1 KC
19 KC
DEM AMPL
DEM CH 4
DEM BF CH 4
0 DB
CHS
SPL SERV
REC AMPL
PROT
0 DB
REC
BP
SPL SERV
-45 DB
IMPEDANCE MATCHING NETWORK

Fig. 7. Block diagram of 4-channel terminal

comparable to that of the commercial equivalents.

2. A modulator and demodulator with good transmission characteristics including prescribed overloading shape stabilized with temperature, excellent carrier balances without adjustments, and uniformity in manufacture.

3. A compact plug-in carrier supply requiring no field adjustments using a rugged stable crystal with regenerative divider circuits and efficient harmonic generator.

4. Line amplifiers having good stability with power, temperature, and tubes; high useful output with low distortion and economy of power. These amplifiers are interchangeable as to "transmitting," "receiving," or "line amplifier" function and mount on a plug-in basis.

5. A ringer-oscillator, used in the orderwire circuit, combining the sending-receiving function in the same two tubes with reliable and interference-free operation.

6. A system alarm covering all equipment in the common path for all channels and operating at a carrier frequency (rather than direct current) to give coverage for the intervening radio links. This employs a high gain, 1-tube circuit making use of a stable reflex feature.

7. Line equalization giving independent flat, slope, and bulge correction under manual control when adjusted in the appropriate sequence.

8. Measuring circuits capable of all essential transmission checks which are simple, stable, and self-contained in the main equipment units.

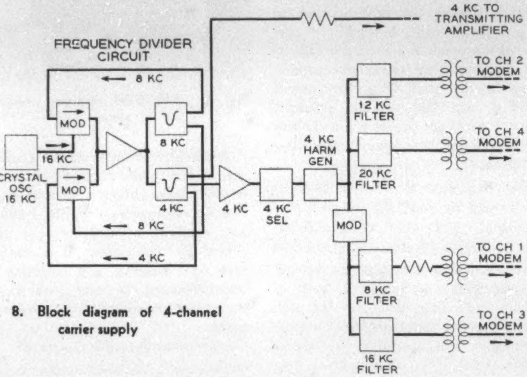

FREQUENCY DIVIDER CIRCUIT
8 KC
MOD
16 KC
CRYSTAL OSC 16 KC
MOD
8 KC
4 KC
8 KC
4 KC
4 KC
4 KC SEL
4 KC HARM GEN
4 KC TO TRANSMITTING AMPLIFIER
12 KC FILTER
TO CH 2 MODEM
20 KC FILTER
TO CH 4 MODEM
MOD
8 KC FILTER
TO CH 1 MODEM
16 KC FILTER
TO CH 3 MODEM

Fig. 8. Block diagram of 4-channel carrier supply

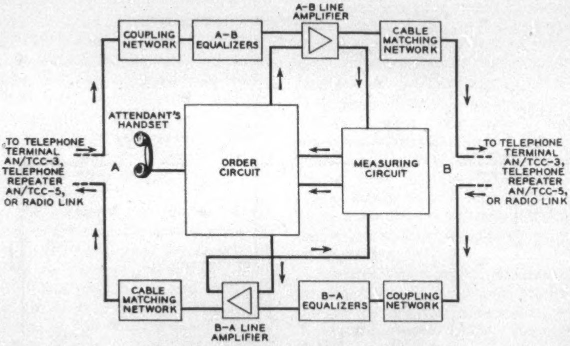

COUPLING NETWORK
A-B EQUALIZERS
A-B LINE AMPLIFIER
CABLE MATCHING NETWORK
ATTENDANT'S HANDSET
TO TELEPHONE TERMINAL AN/TCC-3, TELEPHONE REPEATER AN/TCC-5, OR RADIO LINK
A
ORDER CIRCUIT
MEASURING CIRCUIT
B
TO TELEPHONE TERMINAL AN/TCC-3, TELEPHONE REPEATER AN/TCC-5, OR RADIO LINK
CABLE MATCHING NETWORK
B-A LINE AMPLIFIER
B-A EQUALIZERS
COUPLING NETWORK

Fig. 9. Block diagram of 4-channel repeater

The 12-Channel System

The 12-channel system provides a means of transmitting a band of frequencies 48 kc wide over nonloaded spiral-4 cable equipped with carrier repeaters or over a suitable radio relay system. This band will carry 12 high-grade voice-frequency message channels derived from three modem units of four channels each, the same as those used in *AN/TCC-3* terminals. Any of the 16-kc bands normally occupied by four channels as well as the whole 48-kc band may be used for special services such as high-speed facsimile and data transmission. In addition to the 12 channels, a voice-frequency order circuit is provided for maintenance purposes.

As is seen in Fig. 12, the block schematic of the *AN/TCC-7*, three steps of modulation are used. The initial step is equivalent to that used in the 4-channel system. The second step raises the channels to the 60 to 108-kc band, and the third step translates the channels to their line frequencies. Carrier frequen-

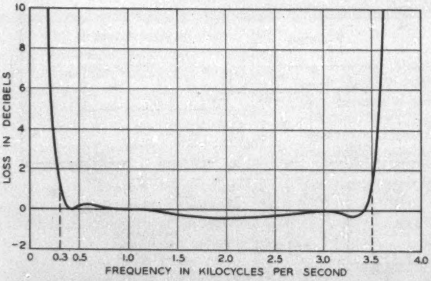


Fig. 10. Channel attenuation characteristic relative to 1,000 cycles per second

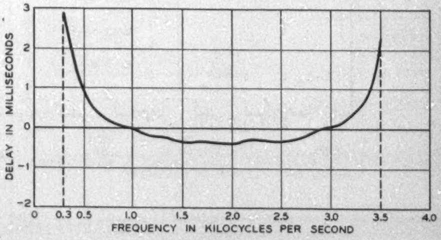


Fig. 11. Channel delay characteristic relative to 1,000 cycles per second

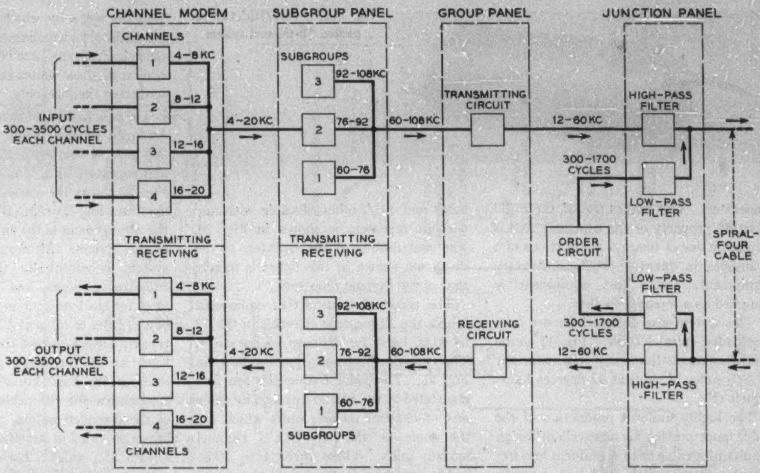

CHANNEL MODEM
SUBGROUP PANEL
GROUP PANEL
JUNCTION PANEL
CHANNELS
1
4-8 KC
2
8-12
3
12-16
4
16-20
INPUT
300-3500 CYCLES
EACH CHANNEL
4-20 KC
SUBGROUPS
3
92-108 KC
2
76-92
1
60-76
60-108 KC
TRANSMITTING CIRCUIT
12-60 KC
HIGH-PASS FILTER
300-1700 CYCLES
LOW-PASS FILTER
ORDER CIRCUIT
SPIRAL-FOUR CABLE
TRANSMITTING
RECEIVING
TRANSMITTING
RECEIVING
1
4-8 KC
2
8-12
3
12-16
4
16-20
OUTPUT
300-3500 CYCLES
EACH CHANNEL
CHANNELS
4-20 KC
3
92-108 KC
2
76-92
1
60-76
SUBGROUPS
60-108 KC
RECEIVING CIRCUIT
12-60 KC
LOW-PASS FILTER
300-1700 CYCLES
HIGH-PASS FILTER

Fig. 12. Block diagram of 12-channel terminal

cies of 8, 12, 16, 20, 56, 72, 88, and 120 kc needed for these steps of modulation are all derived from a 64-kc thermistor-stabilized crystal-controlled oscillator.

The system is made up of telephone terminals *AN-TCC-7*, attended repeaters *AN/TCC-8*, and unattended repeaters *AN/TCC-11*, as shown in Figs. 13, 14, and 15 respectively. A typical cable system layout is shown in Fig. 16. With unattended repeaters spaced at intervals of $5^3/_4$ miles and attended repeaters spaced at 40 miles, a system of this type can be extended to 200 miles. Five of these 200-mile links may be connected in tandem to form a multilink circuit which meets the established performances for long-haul toll circuits.[2–6] As the channel equipment is the same as in the *AN/TCC-3* equipment and as the other parts of the system introduce distortions of less than a few tenths of a db, the channel transmission performance is about the same as that previously shown in Figs. 10 and 11. The transmission of the 48-kc band will be equalized to well within ±2 db for a 200-mile system and the automatic regulation will hold this band to better than ±3 db for at least a ±30-degree temperature change. For greater temperature changes, manual "slope" and "bulge" adjustments are provided to supplement the automatic regulation.

Power for a maximum of three unattended repeaters is supplied from a terminal or in both directions from an attended repeater. Power is applied to the line at a constant current of 100 milliamperes at a nominal voltage of 600 volts. Fig. 17 shows a typical power loop circuit. High- and low-voltage alarm circuits operate relays which automatically remove voltage from the cable to provide protection to personnel.

Complete transmission testing facilities are provided as part of the terminal and attended repeater equipment. These measuring sets plus a portable test set for in-service testing of the unattended repeater provide means for complete system line-up and maintenance. The measuring sets at the attended points include equipment to check the modulation performance of the carrier repeaters, as well as a means for locating faulty unattended repeaters. A unique feature of the portable test set is a transistor oscillator used as a 1,600-cycle-per-second signalling source powered from the rectified output of the lineman's hand-operated

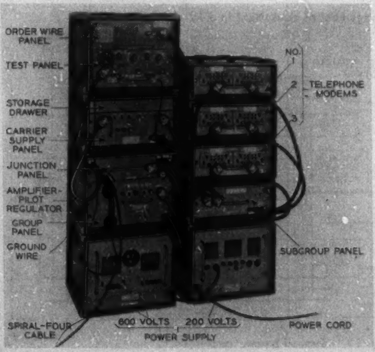

ORDER WIRE PANEL
TEST PANEL
STORAGE DRAWER
CARRIER SUPPLY PANEL
JUNCTION PANEL
AMPLIFIER-PILOT REGULATOR
GROUP PANEL
GROUND WIRE
SPIRAL-FOUR CABLE
NO. 1
2
3
TELEPHONE MODEMS
SUBGROUP PANEL
600 VOLTS
200 VOLTS
POWER SUPPLY
POWER CORD

Fig. 13 (left). AN/TCC-7 terminal: 12-channel system

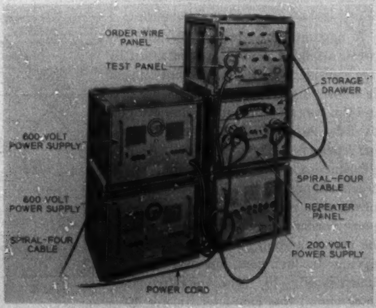

ORDER WIRE PANEL
TEST PANEL
STORAGE DRAWER
600 VOLT POWER SUPPLY
600 VOLT POWER SUPPLY
SPIRAL-FOUR CABLE
SPIRAL-FOUR CABLE
REPEATER PANEL
200 VOLT POWER SUPPLY
POWER CORD

Fig. 14 (right). AN/TCC-8 repeater: 12-channel system

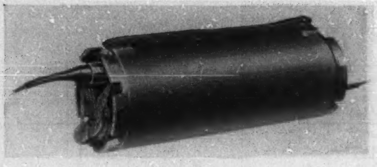


Fig. 15. AN/TCC-11 repeater: 12-channel system

generator. This makes use of the rapid starting property of the transistor and is probably one of the first applications of a transistor in quantity-produced military equipment. Cable test equipment is supplied as a separate facility.

The system was designed for use on a nonloaded spiral-4 cable. Fig. 18 shows the loss-versus-frequency characteristic of $5^{3}/_{4}$ miles of cable at 45 degrees Fahrenheit (F).

The highly uniform production of the cable made possible by an excellent design results in a cable with a uniform loss-frequency characteristic and low capacity unbalances as reported in a companion paper.[7] The uniform loss-frequency characteristic allows the successful operation of a system using only one pilot frequency to automatically regulate the system. The low capacity unbalances result in a working system with noise a near- and intrachannel crosstalk (echo) well within the design objectives.

A line-frequency spectrum of 12 to 60 kc was chosen to obtain a relatively smooth response. Because of the low capacity unbalance between pairs of the cables, a repeater spacing of $5^{3}/_{4}$ miles is permissible when the 12 to 60-kc band is used for both directions of transmission. This results in satisfactory near-end intrachannel crosstalk (echo) of over 25 db for an over-all 200 mile system. Only about a 10-per-cent reduction in repeater spacing was necessary to allow the use of the same frequency band for both directions of transmission. This permitted both terminals to be alike and simplified installation and maintenance.

A maximum spacing of 40 miles between attended points was controlled by two considerations. It did not appear desirable to exceed 600 volts on the line, and this was about the maximum distance an order circuit could be operated without repeaters. The order circuit is made independent of the carrier amplifiers between attended points in order to obtain maximum reliability. The voice-frequency order circuit is by-passed around the carrier amplifier at each unattended repeater by high-pass, low-pass filters as shown in Fig. 19. The voice-frequency order circuit loss characteristic for $40^{1}/_{4}$ miles and $17^{1}/_{4}$ miles of cable with and without repeaters is shown in Fig. 20. The contribution of the repeaters to the distortion shown in this figure is merely that of the by-pass filter sets.

The terminal transmitting equipment spaces the 12 message channels in the 12 to 60-kc band for delivery to the cable. The steps for this translation are shown in Fig. 21. The voice frequencies are first translated to the 4 to 20-kc range by three sets of channel modem units which are the same as the *AN/TCC-3* channel modem units. These three 4 to 20-kc bands are then translated to the 60 to 108-kc range on the upper side bands of carriers of 56, 72, and 88 kc. The 60 to 108-kc band is then translated down to the 12 to 60-kc range by a modulator with a 120-kc carrier. In the receiving direction the steps by which the 12 message channels are demodulated to voice frequencies from the 12 to 60-kc band are the reverse of those shown in Fig. 21 and described in the foregoing.

The equalization of this system has been arranged to give a flat response over a 200-mile system with all cable sections and repeaters at a temperature of 45 F. Fig. 22 shows the various carrier amplifiers and Fig. 23 (A), (B), and (C) show the arrangement of the equalizer networks in the system. All amplification in the system is essentially flat. The basic equalizer provides loss which compensates for the loss-frequency characteristic of $5^{3}/_{4}$ miles of cable at 45 F. The deviation equalizers correct the various deviations at the attended points plus the error of the basic equalizers to exactly compensate for the cable loss. As part of the attended points, adjustable cable loss is provided in steps over a range from 1/4 to $5^{1}/_{4}$ miles. These "building-out networks" allow flexibility in the positioning of attended points in the system. In addition to these equalizers, flat, slope, and bulge adjustments are provided to compensate for the deviation of the cable characteristic and the repeaters from the

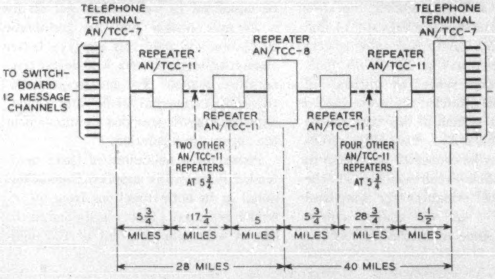

TELEPHONE TERMINAL AN/TCC-7
TO SWITCH-BOARD 12 MESSAGE CHANNELS
REPEATER AN/TCC-11
REPEATER AN/TCC-11
TWO OTHER AN/TCC-11 REPEATERS AT 5 3/4
REPEATER AN/TCC-8
REPEATER AN/TCC-11
FOUR OTHER AN/TCC-11 REPEATERS AT 5 3/4
REPEATER AN/TCC-11
TELEPHONE TERMINAL AN/TCC-7
TO SWITCH-BOARD 12 MESSAGE CHANNELS
5 3/4 MILES
17 1/4 MILES
5 MILES
5 3/4 MILES
28 3/4 MILES
5 1/2 MILES
28 MILES
40 MILES

Fig. 16. Typical arrangement of equipment in a 12-channel system

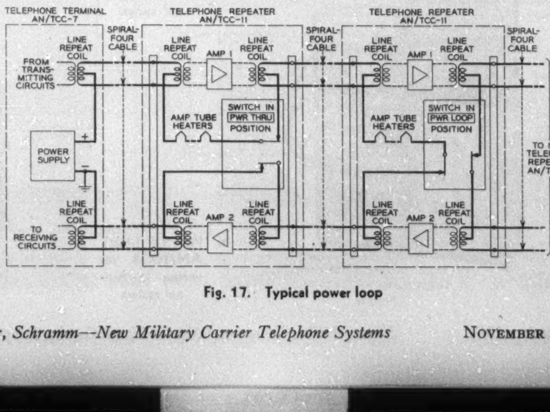

TELEPHONE TERMINAL AN/TCC-7
LINE REPEAT COIL
SPIRAL-FOUR CABLE
FROM TRANS-MITTING CIRCUITS
POWER SUPPLY
+
−
LINE REPEAT COIL
TO RECEIVING CIRCUITS
TELEPHONE REPEATER AN/TCC-11
LINE REPEAT COIL
AMP 1
LINE REPEAT COIL
AMP TUBE HEATERS
SWITCH IN PWR THRU POSITION
LINE REPEAT COIL
AMP 2
LINE REPEAT COIL
SPIRAL-FOUR CABLE
TELEPHONE REPEATER AN/TCC-11
LINE REPEAT COIL
AMP 1
LINE REPEAT COIL
AMP TUBE HEATERS
SWITCH IN PWR LOOP POSITION
LINE REPEAT COIL
AMP 2
LINE REPEAT COIL
SPIRAL-FOUR CABLE
TO NEXT TELEPHONE REPEATER AN/TCC-11

Fig. 17. Typical power loop

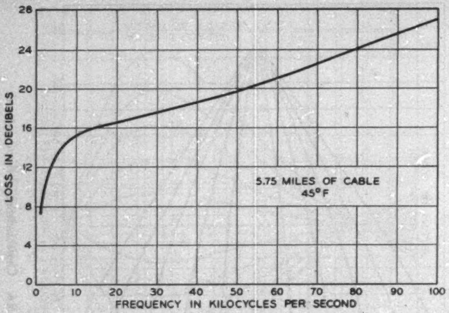


Fig. 18 (left). Loss of 5.75 miles of nonloaded cable (cable CX-1065/G)

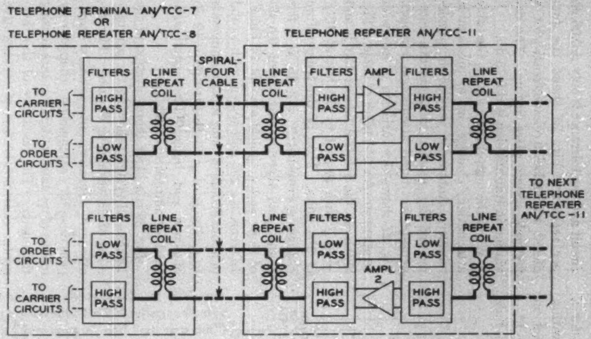

TELEPHONE TERMINAL AN/TCC-7
OR
TELEPHONE REPEATER AN/TCC-8
TELEPHONE REPEATER AN/TCC-11
TO CARRIER CIRCUITS
TO ORDER CIRCUITS
TO ORDER CIRCUITS
TO CARRIER CIRCUITS
FILTERS
HIGH PASS
LOW PASS
LINE REPEAT COIL
SPIRAL-FOUR CABLE
AMPL 1
AMPL 2
TO NEXT TELEPHONE REPEATER AN/TCC-11

Fig. 19 (right). Carrier and order circuits

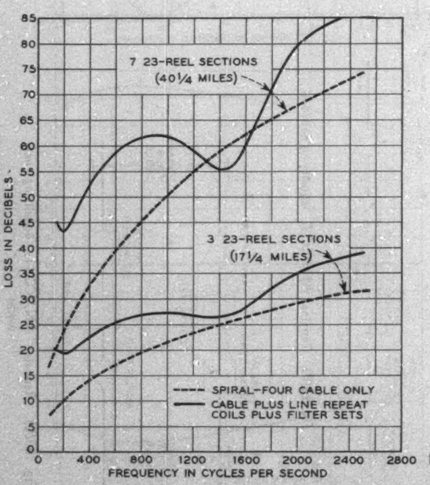


Fig. 20 (left). Order-circuit loss versus frequency

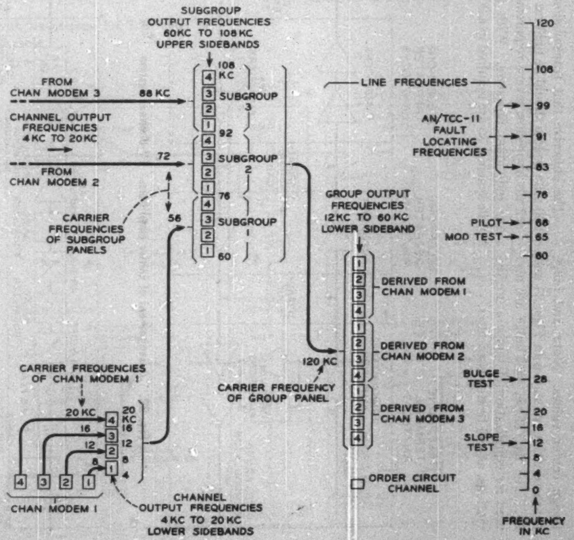

SUBGROUP OUTPUT FREQUENCIES 60KC TO 108KC UPPER SIDEBANDS
FROM CHAN MODEM 3
88 KC
CHANNEL OUTPUT FREQUENCIES 4KC TO 20KC
72
FROM CHAN MODEM 2
CARRIER FREQUENCIES OF SUBGROUP PANELS
56
108 KC
92
76
60
SUBGROUP 3
SUBGROUP 2
SUBGROUP 1
CARRIER FREQUENCIES OF CHAN MODEM 1
20 KC
16
12
8
20 KC
16
12
8
4
CHAN MODEM 1
CHANNEL OUTPUT FREQUENCIES 4 KC TO 20 KC LOWER SIDEBANDS
GROUP OUTPUT FREQUENCIES 12KC TO 60 KC LOWER SIDEBAND
120 KC
CARRIER FREQUENCY OF GROUP PANEL
DERIVED FROM CHAN MODEM 1
DERIVED FROM CHAN MODEM 2
DERIVED FROM CHAN MODEM 3
ORDER CIRCUIT CHANNEL
LINE FREQUENCIES
AN/TCC-11 FAULT LOCATING FREQUENCIES
PILOT
MOD TEST
BULGE TEST
SLOPE TEST
120
108
99
91
83
76
68
65
60
28
20
16
12
8
4
0
FREQUENCY IN KC

Fig. 21 (right). Frequency allocation of 12-channel system

design center values. Slope and bulge characteristics are shown in Figs. 24 and 25. The flat, slope, and bulge are also used for equalization of temperature deviations, as will be discussed later.

The objective that transmission for a 200-mile system should be uniform within ±2 db imposes a severe requirement on the individual repeaters. With a $5^3/_4$-mile spacing, 35 are required in a 200-mile circuit and the tolerance is only ±0.06 db per repeater. Fig. 26 shows the predicted transmission characteristic for a 200-mile system after line-up. The deviation in transmission is ±0.75 db.

When the temperature varies, the cable attenuation changes in a predictable manner. Automatic regulation supplemented by manual adjustments compensate for these attenuation changes. Fig. 27 shows the variation of loss with temperature for $5^3/_4$ miles of cable. Fig. 28 shows the relative loss at 130 F and −40 F with respect to the loss at 45 F of $5^3/_4$ miles. Most of the temperature change, approximately ±2 db, is a flat change and there is only about 0.3 db of slope and a similar amount of bulge. In terms of a 200-mile system with a total of about 800 db of cable loss, this amounts to a change of ±70 db flat and ±10 db of slope and bulge. This emphasizes need for automatic regulation.

Fig. 29 is a block representation of the 12-channel system, showing the arrangement of the regulating circuits. The unattended repeater automatically compensates for attenuation changes by means of an ambient temperature sensing thermistor. The resistance of the thermistor controls the loss of the negative feedback network of the line amplifier. The variations in amplifier gain with thermistor temperature are designed to coincide with the variations in loss of $5^3/_4$ miles of cable over the same temperature range. Thus, if the thermistor temperature follows the cable temperature, correction of the cable loss variations will be obtained. This compensation is only

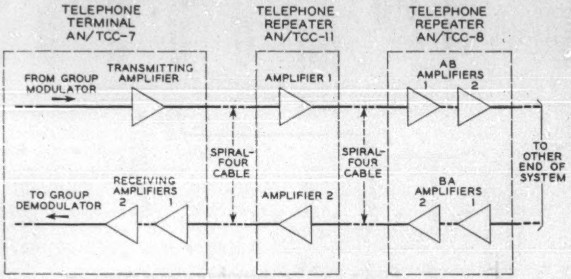

TELEPHONE TERMINAL AN/TCC-7
TELEPHONE REPEATER AN/TCC-11
TELEPHONE REPEATER AN/TCC-8
FROM GROUP MODULATOR
TRANSMITTING AMPLIFIER
AMPLIFIER 1
AB AMPLIFIERS
1
2
SPIRAL-FOUR CABLE
SPIRAL-FOUR CABLE
TO OTHER END OF SYSTEM
RECEIVING AMPLIFIERS
2
1
TO GROUP DEMODULATOR
AMPLIFIER 2
BA AMPLIFIERS
2
1

Fig. 22. Arrangement of carrier amplifiers: 12-channel system

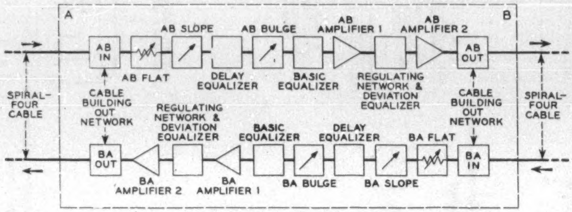

A
B
AB SLOPE
AB BULGE
AB AMPLIFIER 1
AB AMPLIFIER 2
AB IN
AB FLAT
DELAY EQUALIZER
BASIC EQUALIZER
REGULATING NETWORK & DEVIATION EQUALIZER
AB OUT
SPIRAL-FOUR CABLE
CABLE BUILDING OUT NETWORK
REGULATING NETWORK & DEVIATION EQUALIZER
BASIC EQUALIZER
DELAY EQUALIZER
BA FLAT
CABLE BUILDING OUT NETWORK
SPIRAL-FOUR CABLE
BA OUT
BA IN
BA AMPLIFIER 2
BA AMPLIFIER 1
BA BULGE
BA SLOPE

Fig. 23(A). Repeater AN/TCC-8: equalizing networks

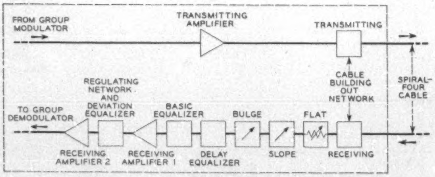

FROM GROUP MODULATOR
TRANSMITTING AMPLIFIER
TRANSMITTING
CABLE BUILDING OUT NETWORK
SPIRAL-FOUR CABLE
REGULATING NETWORK AND DEVIATION EQUALIZER
TO GROUP DEMODULATOR
BASIC EQUALIZER
BULGE
FLAT
RECEIVING AMPLIFIER 2
RECEIVING AMPLIFIER 1
DELAY EQUALIZER
SLOPE
RECEIVING

Fig. 23(B). Terminal AN/TCC-7: equalizing networks

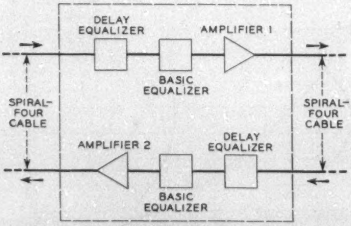

DELAY EQUALIZER
AMPLIFIER 1
BASIC EQUALIZER
SPIRAL-FOUR CABLE
SPIRAL-FOUR CABLE
AMPLIFIER 2
DELAY EQUALIZER
BASIC EQUALIZER

Fig. 23(C). Repeater AN/TCC-11: equalizing networks

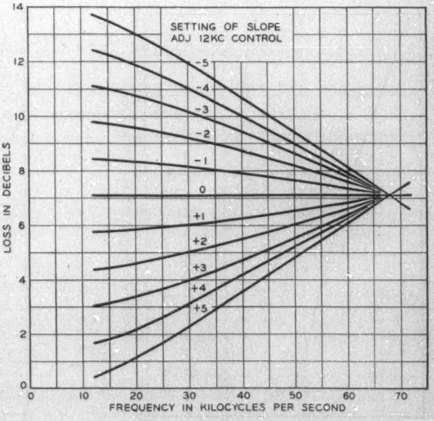
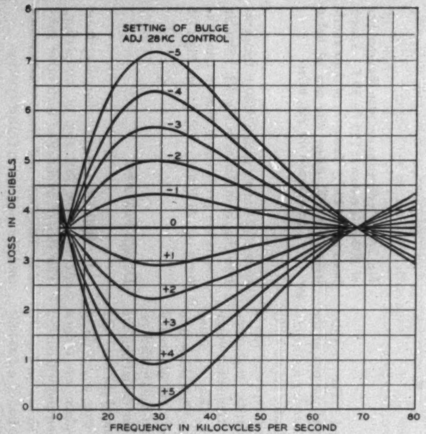
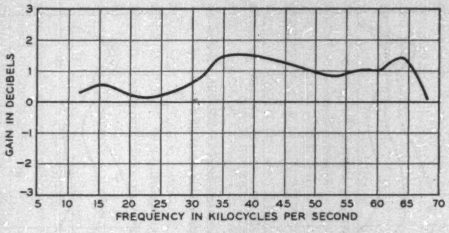


Fig. 24. Characteristics of slope network





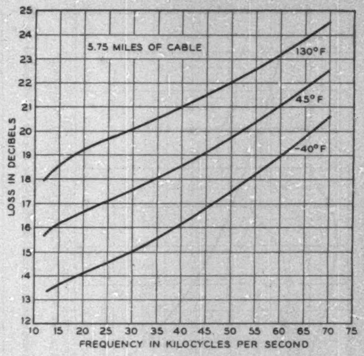


Fig. 25 (left). Characteristics of bulge network

Fig. 26 (above). Predicted transmission characteristic of 200-mile system at 45 F. (One step of slope equalization for 200 miles; no bulge equalization; transmission between "trans amp in" and "rec amp out")

Fig. 27 (right). Variations of loss with temperature for 5.75 miles of nonloaded cable

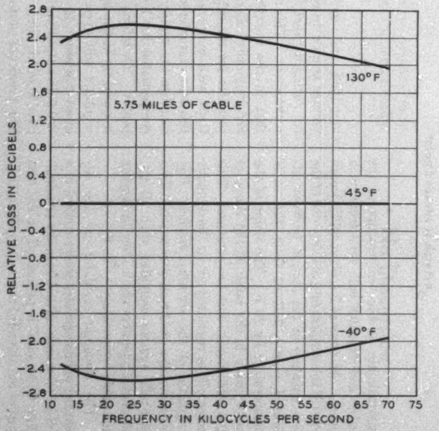


Fig. 28 (left). Relative cable loss at different temperatures

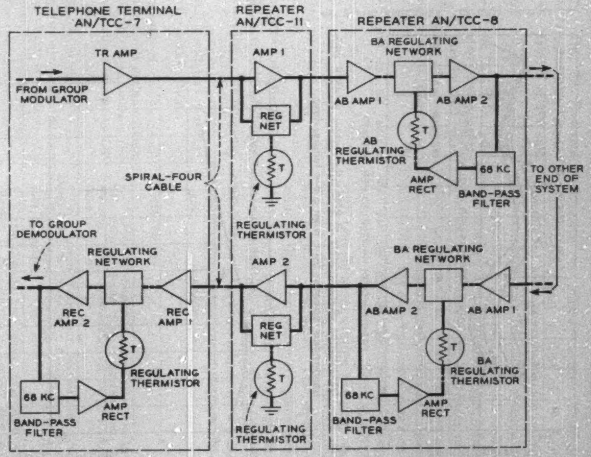

TELEPHONE TERMINAL
AN/TCC-7
REPEATER
AN/TCC-11
REPEATER AN/TCC-8
TR AMP
FROM GROUP
MODULATOR
AMP 1
REG
NET
REGULATING
THERMISTOR
BA REGULATING
NETWORK
AB AMP 1
AB AMP 2
AB
REGULATING
THERMISTOR
AMP
RECT
68 KC
BAND-PASS
FILTER
TO OTHER
END OF
SYSTEM
SPIRAL-FOUR
CABLE
TO GROUP
DEMODULATOR
REGULATING
NETWORK
REC
AMP 2
REC
AMP 1
REGULATING
THERMISTOR
68 KC
BAND-PASS
FILTER
AMP
RECT
AMP 2
REG
NET
REGULATING
THERMISTOR
BA REGULATING
NETWORK
AB AMP 2
AB AMP 1
BA
REGULATING
THERMISTOR
68 KC
BAND-PASS
FILTER
AMP
RECT
T

Fig. 29 (right). Twelve-channel system: arrangement of regulating circuits

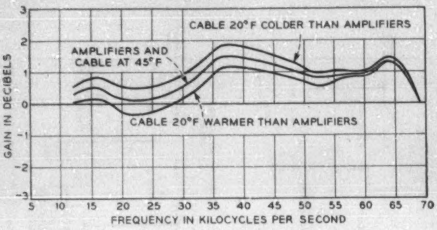


Fig. 30. Transmission characteristics of 200-mile system at 45 F with unattended repeaters 20 F from cable temperature

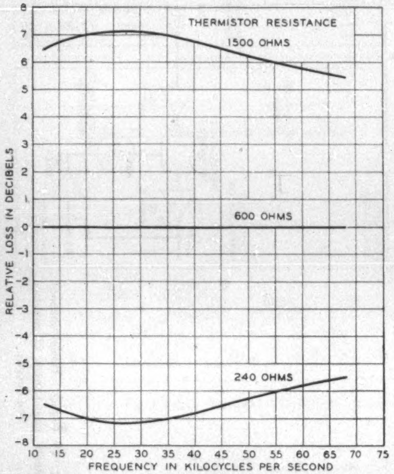


Fig. 31. Relative loss versus thermistor resistance of regulating network at attended point

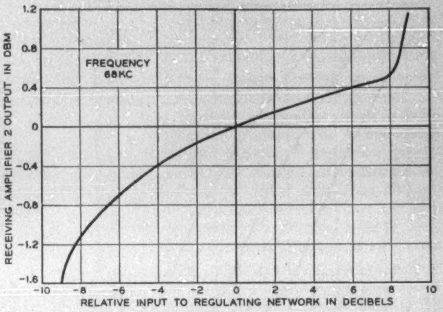


Fig. 32. Output versus input of regulator circuit at attended point

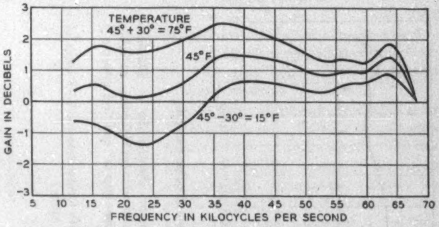


Fig. 33. Predicted transmission characteristics of 200-mile system at 45 F+30 F. System lined up at 45 F

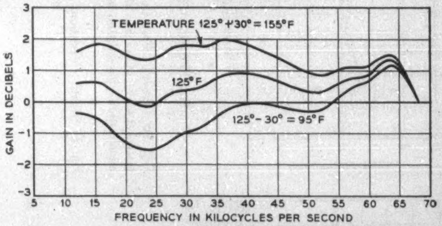


Fig. 34. Predicted transmission characteristics of 200-mile system at 125 F±30 F. System lined up at 125 F. Two steps of slope equalization for 200 miles; one step of bulge equalization for 200 miles

approximate as the temperature of the cable and repeater will vary from each other. An error of 2.4 db at 68 kc would exist if the temperature of every one of the six repeaters in the section between attended points differed from the cable temperature by 20 F. The shape of the error curve is the same as the shape of the cable change with temperature. This error will be regulated out by the dynamic regulator at the attended station. Fig. 30 shows the predicted performance for a 200-mile system assuming that all unattended repeaters are leading or lagging the cable temperature by 20 F.

The dynamic regulator at the attended points operates under control of a 68-kc pilot transmitted from the terminal. The 68-kc pilot is picked off at the output of the attended station amplifier, then amplified and rectified. The rectified voltage is used to control the loss of the regulating network by varying a thermistor element. This regulator network like that in the unattended repeater has a loss-frequency characteristic that compensates for the cable attenuation variations with temperature. Fig. 31 shows this loss characteristic. This shape also matches the temperature deviations not compensated for by the unattended repeater. Therefore, these deviations are mopped up at the attended points. The regulator will operate over a ±5 db of input variation to produce less than a ±0.5 db output variation and will control over a wider range as shown on Fig. 32. The predicted performance of a 200-mile system operating with a ±30 F temperature change is shown in Fig. 33. Similar results are predicted at line-up temperatures of 125 F and −30 F; see Figs. 34 and 35.

Design Features

1. Both the 4- and 12-channel systems use the same first step of modulation. This step of modulation with other suitable subgroup and group modulations could be used to form other systems.

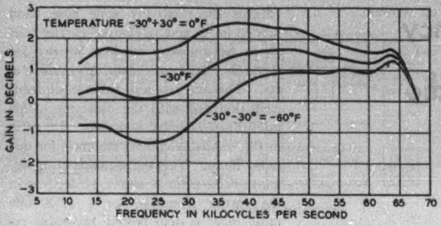


Fig. 35. Predicted transmission characteristics of 200-mile system at −30 F+30 F. System lined up at −30 F. One step bulge equalization for 200 miles; no slope equalization

2. The combination of the first and second steps of modulation in the *AN/TCC-7* terminal form a 12-channel group of 4-kc spaced channels similar to that which is used in the long-haul Bell System *J*, *K*, and *L* systems. This 60 to 108-kc group is internationally recognized and could be a patch point between military and both American and European long-haul carrier telephone systems.

3. Built-in test features in both systems permit all normal maintenance while the channels are in service.

4. In the 4-channel system, all normal regulation can be accomplished by adjustment of a single flat gain control. By adjusting transmission on the order circuit, it is possible to control the transmission of the carrier channels.

5. In the 12-channel system, the regulation is precise and automatic for as much as ±30 F from the line-up temperature. Compensation or realignment for further temperature change can be made by measurement of test pilots introduced between channels during regular operation of all 12 channels. Routine gain checks and even filament activity checks of the unattended repeaters can be made during all service conditions.

6. All measurement or maintenance of carrier unbalances is unnecessary as the primary modulator elements are selected at the time of manufacture and stabilized by added resistors and thermistors.

7. Other thermistors stabilize the loss and the limiting characteristics of the primary modulators over the entire temperature range.

8. Stable crystal oscillators generate the base frequency in both systems.

a. The 16-kc oscillator in the 4-channel terminal is sufficiently stable so that without field adjustment the difference between the carriers of any two terminals will not exceed ±2 cycles.

b. The 64-kc oscillator in the 12-channel terminal is temperature-controlled by a thermistor. This oscillator is stable to ±3 parts in 10^6; however, in system operation all frequencies at the two terminals are derived from the control terminal. No field adjustment of frequency is necessary.

9. Both systems are designed with excess bulge and slope equalizer range so that other transmission media may be used.

10. Use of the latest techniques in plastic sealing have resulted in many new miniaturized components, all complying with the latest military specifications.

Conclusions

Two carrier telephone systems have been developed to meet the severe design objectives of military usage. Channel performance equal to that of the best commercial systems, suitable for multilink operation with telephone, telegraph, or facsimile has been attained with highly miniaturized and ruggedized equipment. The frequency selectivity and frequency stability required to permit efficient use of the frequency spectrum has been achieved by the development of new components and by the application of group modulation techniques. The stability with time and temperature has been obtained by the liberal use of thermistors and negative feedback. The development of these new carrier telephone systems, together with the associated development of new spiral-4 cable and new linear radio sets, forms a basis for a comprehensive military communication network.

References

1. A New Ultrahigh-Frequency Multichannel Military Radio Relays System, **John G. Nordahl.** *AIEE Transactions*, vol, 73, pt. I, Nov. 1954, pp. 526–31.

2. A Carrier Telephone System for Toll Cables, **C. W. Green, E. I. Green.** *Bell System Technical Journal*, New York, N. Y., vol. 17, Jan. 1938, pp. 80–105.

3. Cable Carrier Telephone Terminal, **R. W. Chesnut, L. M. Ilgenfritz, A. Kenner.** *Ibid.*, pp. 106–24.

4. A Twelve-Channel System for Open-Wire Lines, **H. A. Affel, B. W. Kendall.** *Ibid.*, vol. 18, Jan. 1939, pp. 119–42.

5. Frequency Division Techniques for a Coaxial Cable Network, **R. E. Crane, J. T. Dixon, G. H. Huber.** *AIEE Transactions*, vol. 66, 1947, pp. 1451–59.

6. An Improved Cable Carrier System, **H. S. Black, F. A. Brooks, A. J. Wier, I. G. Wilson.** *Ibid.*, pp. 741–46.

7. A New Military Cable Design for Military Carrier Telephone Systems, **H. F. X. Kingsley.** *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 506–09.

◆

No Discussion

A New Ultrahigh-Frequency Multichannel Military Radio Relay System

JOHN G. NORDAHL
NONMEMBER AIEE

MANY items must be studied in designing a new radio set, and a large amount of information and data must be obtained which can therefore be exhibited. However, the practical confines of this paper make it possible to cover only in a general way the major items which may be of interest to other development engineers. Accordingly, this paper will first outline what the new military radio relay equipment does and what it consists of; then briefly describe the circuits of the major units; after which much of the text will cover the important and difficult problems of distortion and noise which so directly affect multichannel systems performance; and will close with a few remarks concerning some of the systems' problems and their solutions. In a companion paper V. I. Cruser[1] describes the many equipment features of this radio set.

A paper by Ribe and Brown[2] discusses the military requirements for the *AN/TRC-24* radio set as a system, and notes the considerations involved in establishing the specification requirements. It was up to those in the development laboratory, then, to attain those requirements in actual equipment. There was no earlier radio set upon which the design could be directly based. It was necessary to start with a fresh viewpoint and work up to a complete radio system, from power supply, through the radio equipment, to the antennas. Many mechanical problems were intertwined with electrical requirements and each electrical requirement was affected by others. Accordingly, the project was necessarily set up on a project engineer basis. With the co-operation of Signal Corps engineers, it was possible to decide early in the development period on a general course of action which ultimately led to the present design.

Description

Radio Set

The *AN/TRC-24* radio set consists of a radio transmitter, a radio receiver, antenna system, power supply, and accessory units which provide radio transmission and reception on a large number of frequency-modulated radio channels in two frequency bands: 100 to 225 megacycles (mc) and 225 to 400 mc. The two frequency bands are completely covered by the use of continuously adjustable plug-in tuning units. Each radio channel will carry up to 12 4-kc carrier telephone channels plus a control or order wire channel. With suitable terminal apparatus the relatively wide base band needed for 4- and 12-channel carrier telephone transmission can also be used for facsimile, teletype, data transmission, etc. The antenna and radio equipment were designed with the objective of attaining a 50-decibel signal-to-noise ratio in the carrier telephone channels for one radio jump. This corresponds to a signal-to-noise ratio of about 34 decibels in the carrier channels, for a 1,000-mile radio relay system having favorable repeater locations, and equipment spacings of about 25 to 30 miles.

The major packages developed for this radio set and the interconnection of the operating units are shown in Fig. 1; the equipment is shown in its simplest form. Actually, a complete repeater or terminal station will have spare equipment, and a repeater station will also employ additional equipment for transmission and reception in two directions, as noted in the Ribe and Brown paper.[2]

Radio Transmitter

The circuit problems were mainly encountered in the more complicated units: the transmitter and the receiver. A block diagram of the transmitter is shown in Fig. 2. The radio transmitter starts with a clean-up filter and level control in the input circuit block, and a base-band amplifier, the output of which frequency-modulates an oscillator by means of reactance tubes. The modulated output then passes through a series of buffer amplifier stages to driver amplifier, multiplier, and power amplifier stages and to the antenna. The transmitter delivers over 100 watts of radio-frequency (r-f) power. However, it is possible to reduce the power to as low as 10 watts by means of a control on the power supply panel.

The oscillator frequency was made as high as practicable to reduce the number

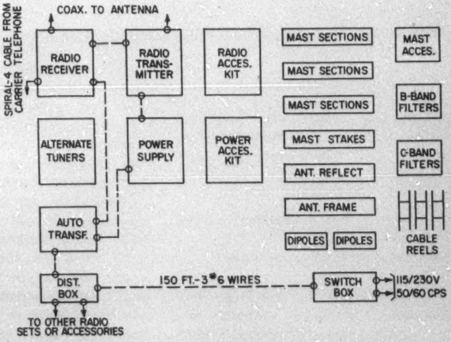


Fig. 1. Major packages developed for radio set AN/TRC-24 and interconnection of the operating units

---

Paper 54-405, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 18, 1954; made available for printing September 1, 1954.

John G. Nordahl is with the Bell Telephone Laboratories, Inc., Murray Hill, N. J.

The development of the *AN/TRC-24* radio set has been a large but interesting project. It could not have reached a successful conclusion without the constant help and encouragement of the Signal Corps engineering personnel during the 2-year development period, as well as of the Bell Telephone Laboratories personnel who gave unstintingly of their time and energy to develop and design this complete radio system.

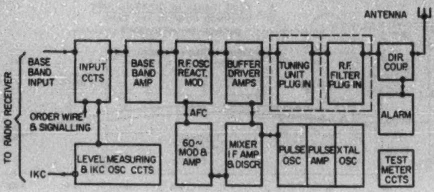

ANTENNA
TO RADIO RECEIVER
BASE BAND INPUT
ORDER WIRE & SIGNALLING
IKC
INPUT CCTS
BASE BAND AMP
R.F. OSC REACT. MOD
BUFFER DRIVER AMPS
TUNING UNIT PLUG IN
R.F. FILTER PLUG IN
DIR COUP.
AFC
LEVEL MEASURING & IKC OSC CCTS
60~ MOD & AMP
MIXER I F AMP & DISCR
PULSE OSC
PULSE AMP
XTAL OSC
ALARM
TEST METER CCTS

Fig. 2. Block diagram of radio transmitter

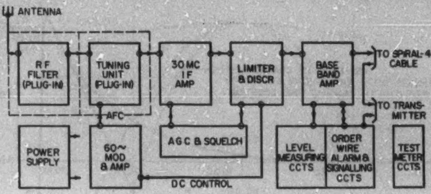

ANTENNA
R F FILTER (PLUG-IN)
TUNING UNIT (PLUG-IN)
30 MC I F AMP
LIMITER & DISCR
BASE BAND AMP
TO SPIRAL-4 CABLE
TO TRANSMITTER
AFC
POWER SUPPLY
60~ MOD & AMP
A G C & SQUELCH
DC CONTROL
LEVEL MEASURING CCTS
ORDER WIRE ALARM & SIGNALLING CCTS
TEST METER CCTS

Fig. 3. Block diagram of radio receiver

of spurious radiations. It and the buffer amplifiers operate over a frequency range of 50 to $112^{1}/_{2}$ mc, so by doubling once or twice in the later amplifiers each of the two frequency bands may be covered. Other frequency bands may be covered, of course, by suitable design of the tuning units. This also means that the frequency deviation and channel spacing vary with the frequency band or tuning unit used.

The transmitter oscillator is continuously tunable, but once adjusted its frequency is maintained by means of an automatic-frequency-control capacitor on the shaft of a geared-down motor which receives its control signal from the circuits indicated in the lower right of the block diagram in Fig. 2. Here a quartz crystal oscillator operating into the pulse amplifier and pulsed oscillator produces a "picket fence" of output frequencies. Each picket is spaced 1/2 mc from its neighbors. The oscillator frequency is beat against one of the picket frequencies to produce a $10^{1}/_{8}$ mc signal. This signal is passed through an intermediate-frequency amplifier, a discriminator, and a 60-cycle modulator-amplifier for control of the automatic-frequency-control motor. By operating the pulsed oscillator above and below the oscillator frequency it is possible to control the oscillator frequency at 1/4 mc intervals even though the pickets are 1/2 mc apart. The transmitter also contains order wire, metering, and alarm circuits which provide for control, adjustment, and surveillance of the operation during service as well as at the test bench.

Table I. Systems Data

	B Band	C Band
Frequency range, mc	100 to 225	225 to 400
Channel spacing, mc	0.5	1.0
Number of channels	250	175
Frequency deviation, kc	±240	±400
Modulation	Modified frequency modulation	
Base-band channels	1 speech and 4 or 12 carrier telephone	
Base-band frequency range, cycles per second	250 to 68,000	
Signaling frequency, cycles per second	1,600	
Line-up frequencies, cycles per second	1,000 and 68,000	

The channel spacing, number of channels, and frequency deviation determined by this transmitter design are indicated in Table I, together with some other systems data which will be useful later in this paper.

Radio Receiver

A block diagram of the receiver is shown in Fig. 3. It is a single conversion superheterodyne having plug-in tuners for covering the frequency bands mentioned previously, a 6-stage intermediate-frequency amplifier, a 2-stage limiter, a discriminator, a base-band amplifier panel, and an automatic frequency control amplifier and motor circuit to keep the receiver locked onto the incoming r-f signal. The receiver also includes orderwire amplifiers, metering and alarm circuits, and circuits for checking the operation of the radio receiver while in use as well as on the test bench.

Much effort was expended on the receiver r-f tuners to reach optimum noise figure, gain, selectivity, and oscillator stability. The oscillator for the 100- to 225-mc band is ganged to and uses the same type of tuning elements as the r-f circuits. However, the oscillator for the 225- to 400-mc band is of the butterfly type and separately tuned from the r-f stages. This design was chosen so that emphasis could be placed on obtaining the high oscillator stability needed to meet the test and operating requirements set up by the Signal Corps. As an aid in setting the receiver to frequency, a "calibrator" circuit is provided which produces signals at 11-mc intervals throughout the bands during dial adjustment or calibrating procedures. More will be said later about the intermediate-frequency amplifier and the discriminator. The limiter, together with the discriminator, have been made wide band as compared to the intermediate-frequency in order to improve adjacent channel interference characteristics.

Interchannel Crosstalk

It was realized from the first that the major development effort would have to be concentrated on reducing distortion in order to meet the interchannel crosstalk objectives. The distortion requirements were such that no available measuring equipment was good enough from this standpoint. In fact, the distortion through a system of one transmitter and one receiver had to be approximately an order of magnitude lower than anything we had done before.

Measurement

The technique employed for determining whether or not the system cross-

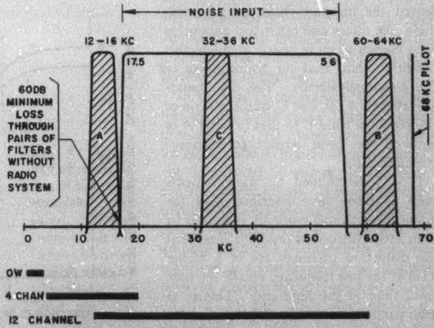


Fig. 4. Passbands of filters employed in the measurement of system interchannel crosstalk ratio, in relation to pilot, orderwire, and carrier frequencies

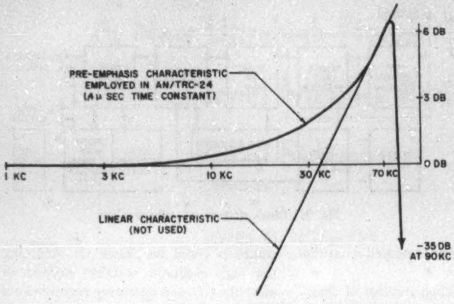


Fig. 5. Pre-emphasis characteristic of radio transmitter

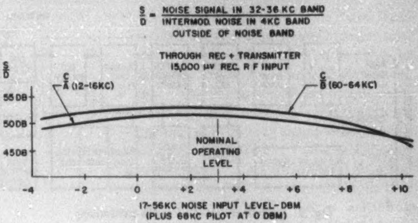


Fig. 6. Interchannel crosstalk ratio versus random noise input level, at a high r-f receiver input

talk objective was attained was suggested by the Signal Corps and it was later found that it had also been employed by other groups in the Bell Telephone Laboratories for certain microwave equipments. Wide-band noise is used to simulate transmission of many modulated carrier telephone channels. A noise generator is connected through a wide but sharply defined band-pass filter to the transmitter input, together with the required 0 level 68-kc pilot frequency normally in use. Then the noise level in a 4-kc band within the noise band and the noise level in a 4-kc band just outside the noise band are measured. The ratio of these two indicate the signal-to-interchannel distortion, or interchannel crosstalk ratio. A wide-band (17 to 56 kc) noise level of 3 decibels above 1 milliwatt at the input to the transmitter was decided upon after a detailed and lengthy study of data on military talker levels, average channel use, peak to average ratios, and other factors.

Fig. 4 shows the filter passbands used in this test, and how they compare with the frequencies covered by the 4- and 12-channel carrier telephone equipments. The transmission of the 12-channel carrier telephone base-band signals presented the greatest challenge. Accordingly, the filters and other test apparatus were chosen to more nearly simulate this condition than that for transmission of the 4-channel signals. The 4-kc bands are lettered *A*, *B*, and *C*, and the ratio of the power in band *C* to that in band *A* or band *B* has been termed interchannel crosstalk ratio.

It was determined by measurements such as those just described that the major interchannel crosstalk is caused by even order distortion products, and that the distortion products are of a higher level at the low frequency end of the base band than at the high end. This is at least partially caused by the fact that the relatively high level pilot frequency is located at the high end of the band (68 kc) and thus the second-order products of this pilot beating with the noise fall in the low end of the band and are of higher level than the third- and fifth-order products which fall mainly in the high end of the band.

EFFECT ON PRE-EMPHASIS CHARACTERISTIC

Interchannel crosstalk produced by distortion in the equipment is only one of the two factors with which to be concerned, the second factor being fluctuation noise introduced at the front end of the receiver. It is well known that the fluctuation noise level in a frequency-modulated receiver increases with frequency in the audio or base band. The first thoughts concerning this system, therefore, were that to obtain a uniform signal-to-noise ratio in each of the 12 carrier telephone channels the modulation should increase linearly with the frequency of the carrier telephone channel. However, when the interchannel crosstalk generated within a system is taken into consideration and the crosstalk is of a somewhat higher level in the low end of the base band, it is necessary to use pre-emphasis and de-emphasis characteristics which are not linear. Data taken under conditions corresponding to operating conditions, i.e., with different received r-f signal strengths, indicated that the best compromise pre-emphasis characteristic for the transmitter is that shown in Fig. 5. This characteristic was obtained by a simple resistance-capacitance network. The de-emphasis characteristic employed in the receiver is the negative of this to produce a flat transmitter-plus-receiver fidelity curve from about 500 cycles per second to 68 kc. Production receivers and transmitters are held to a nominal 0.3 db of these curves by very close control of components, so that any transmitter may be used with any receiver and still achieve an essentially flat response for the radio system as a whole.

The compromise just noted comes about because the base-band signal level will vary (causing more or less interchannel crosstalk) and the r-f signal will vary (causing more or less fluctuation noise) under different operating situations. The emphasis characteristic chosen will produce equal signal-to-noise ratios in the low and high carrier telephone channels under nominal channel loading conditions and at a r-f input level of about 50 microvolts across the 50-ohm receiver antenna connector.

EFFECT ON OPERATING PROCEDURES

A typical curve of interchannel crosstalk ratio versus wide-band noise input level is shown in Fig. 6. For these data,

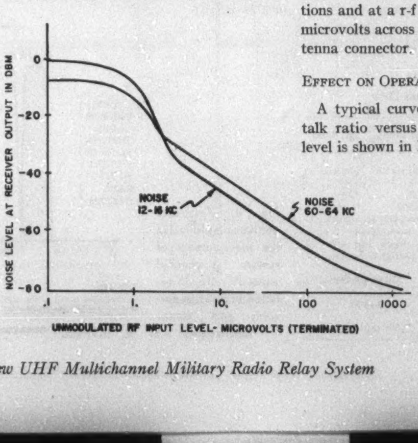


Fig. 7. Receiver quieting characteristic for two carrier telephone channels

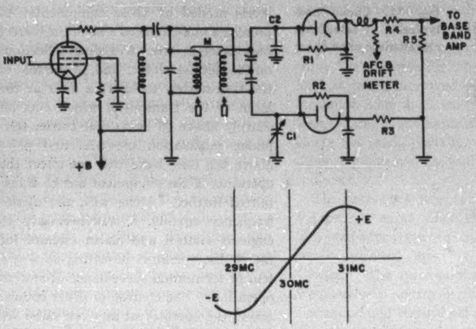


Fig. 8. Simplified schematic circuit of receiver discriminator

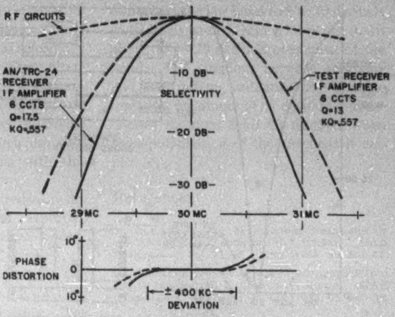


Fig. 10. Receiver intermediate-frequency selectivity and phase distortion characteristics

the pre- and de-emphasis characteristics mentioned previously were employed, and the receiver was operated at a high r-f input to eliminate fluctuation noise from the measurement. Note that the interchannel crosstalk ratio, shown as an S/D ordinate, is fairly constant with moderate changes in the base-band input level from the 3 decibels above 1 milliwatt nominal value. This means that the crosstalk level varies nearly linearly with the base-band input level.

A typical receiver quieting curve is shown in Fig. 7. Fluctuation noise is plotted as the ordinate and r-f input as the abscissa. Note that at the r-f input level for good systems performance which was just mentioned there is a fairly large fading range before the system "breaks" at the 2-microvolt point. Also note the relative level of the fluctuation noise in the 12 to 16-kc and 60- to 64-kc bands.

The two sets of data (Figs. 6 and 7) show that, when one jump in a radio link must be longer than desired due to terrain or other conditions, the lower received r-f signal strength will produce a higher than nominal fluctuation noise, but the signal-to-noise ratio in the carrier telephone channels may be improved by an increase in the base-band input level to the transmitter. The objective is to make the crosstalk noise about equal to the fluctuation noise, within the capabilities of the equipment.

Reduction of Crosstalk

The reduction in distortion which produces interchannel crosstalk was achieved not by the use of any fundamentally new tricks or circuitry but by careful application and extension of known techniques. For an equipment like this which must be safe, stable and permanent, there is a feeling of assurance when the design is based on familiar ground. As examples of this, a closer look will be given to the receiver discriminator and the transmitter modulator, the two major sources of distortion.

A schematic of the receiver discriminator is shown in Fig. 8. Note that the circuit is essentially the familiar Foster-Seely type but has a number of additional components. At the center of a discriminator curve the output can be expected to be relatively linear with input if both sides of the circuit are balanced. A slight unbalance may cause the slope above and to the right of the center to be different than the slope of the characteristic below and to the left. This means distortion. It was found necessary to pay extremely careful attention to balancing: by shunting the diodes (*R*1 and *R*2) in order to make the circuit less affected by variations in these elements; by balancing both sides of the audio output (*R*3, *R*4, and *R*5) instead of just grounding one side of the discriminator output circuit as is normally done; and by adding a variable capacitor *C*1 and a fixed capacitor *C*2 so that the coupling to ground from each side of the discriminator secondary could be balanced. In these circuits, as well as other critical circuits, elements with tolerances no greater than 1 per cent and stable with changes in temperature and time have been selected for long-time, trouble-free operation.

In the case of the transmitter modulator, the use of reactance tubes was noted previously. These tubes are used only for modulation so that their operation will not be affected by automatic-frequency-control considerations. A simplified schematic of the oscillator-reactance tube circuit is shown in Fig. 9. This is a not unusual push-pull circuit which has been improved in a number of ways. One has been to choose the grid-phasing circuit *C*1, *C*2, *R*1 and other elements so as to produce as nearly as possible a 90-

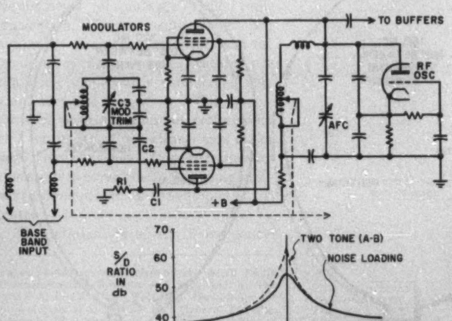


Fig. 9. Simplified schematic circuit of transmitter modulator, and curve showing improvement in interchannel crosstalk ratio by adjustment of "mod trim" capacitor

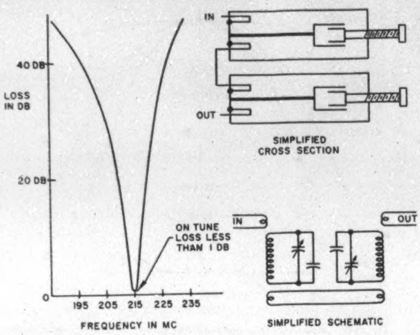


Fig. 11. Circuit and typical selectivity characteristic of r-f filters

degree phase shift between the plate voltage and the grid voltage; a second has been to employ a small variable "mod trim" capacitor $C3$, to accurately tune the grid circuit; the third has been to provide some cathode feedback in both the oscillator and the reactance tubes; and the fourth to insure balance in the audio input circuits. The "mod trim" capacitor is adjusted during the field setup operation to produce maximum modulation sensitivity which has been found to correspond with minimum interchannel distortion. This also is shown in Fig. 9 for both noise and 2-tone modulation.

It was possible to operate on other portions of the receiver and transmitter to reduce their distortion sufficiently below that produced by the receiver discriminator and the transmitter modulator. For instance, the receiver intermediate-frequency circuits, if improperly designed, could produce phase distortion. Studies indicated, however, that this could be reduced to a low value, and still meet the selectivity requirements, by use of undercoupled circuits in the intermediate-frequency networks. Fig. 10 shows the selectivity and phase distortion for two values of circuit Q, one as used in the *AN/TRC-24* receiver to meet selectivity requirements, and the other as used in laboratory test receivers to reduce the intermediate-frequency distortion so that it could be ascertained that discriminator or modulator distortion was being measured during the studies of those circuits. Other tuned circuits in the receiver and in the transmitter were designed in a similar manner to cause little distortion and yet produce sufficient selectivity for the job each particular circuit had to do.

Another probable source of distortion is the audio or base-band amplifier used in both the transmitter and the receiver. These are 3-tube amplifiers similar in general design to those employed in the *AN/TCC* carrier telephone equipments. These amplifiers follow feedback amplifier design techniques and are about 8 to 10 decibels better from the distortion standpoint than the discriminator and modulator circuits. They therefore do not materially contribute to the over-all distortion.

Systems Aspects

There were many systems aspects to be considered during the development. In addition to the power supply problems, which will not be considered here, there were those base-band problems concerning interconnection with carrier telephone equipments and those problems having to do with the r-f parts of the system.

BASE-BAND PROBLEMS

Interconnection with the *AN/TCC* equipments required the radio set to operate at input and output base-band levels needed by those equipments; to present a well-balanced load of the correct impedance in order to reduce reflections on the spiral-4 cables; and, in the case of the transmitter, to insert a filter at the input of the transmitter which cuts off sharply above 68 kc so that carrier telephone modulation products and noise above the base band will not affect the operation of the equipment nor be transmitted further. Along with the carrier frequency circuits, it was necessary to engineer control and alarm circuits for use by the operators in setting up a system or for normal surveillance of system operation. The control, or order circuit, allows the operator at any one radio set to talk and listen both ways on the radio and carrier circuit. A 1,600-cycle ringing circuit is provided so that he may call other stations along the line. A 1,000-cycle oscillator and tuned metering circuits are also provided in the equipment for adjustment of levels up and down the radio system prior to connection of the carrier telephone equipments. A 68-kc filter, together with the metering circuits, provides for adjustment of levels and continuous surveillance of the circuit when the 12-channel carrier telephone system is operating through this radio set.

R-F SYSTEMS IMPROVEMENTS

Considerable effort was applied to improve this equipment from the r-f systems standpoint. It has been noted previously that the transmitter oscillator frequency was chosen as high as possible in order to reduce the number of spurious frequencies (harmonics) radiated and that attention had been given to selectivity in the design of both the receiver and the

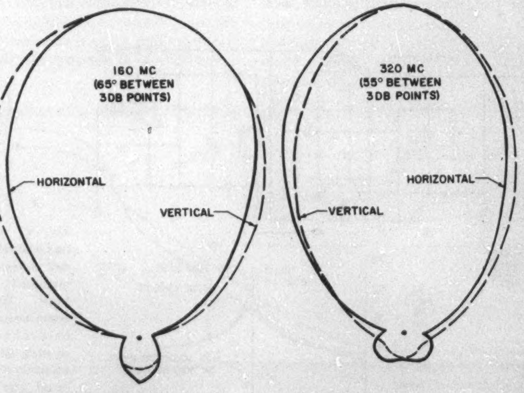


Fig. 12. Reflector-type antenna directivity at two frequencies

transmitter. One of the more unusual problems, however, concerned the need for placing a receiving and transmitting antenna on the same mast. A high transmitter signal then appears at the receiver through the coupling between antennas to cause overloading and the generation of spurious frequencies in the early stages of the receiver. Much attention was given during the design of the antenna to reducing coupling between the two antennas on the same mast and still have a practical mechanical structure. However, the major improvement in circuit operation was obtained by the design and use of tunable cavity filters for insertion in both the transmitter and the receiver antenna circuits. This solution will be recognized by those familiar with mobile radio interference problems. The internal arrangement of these filters, each one covering about 20 to 30 per cent of its mid-band frequency, as well as a typical selectivity curve, are shown in Fig. 11. These filters have so much improved r-f systems performance from both the receiving and transmitting standpoints that the Signal Corps has made plans for their use with other equipments as well as with this radio set.

The antenna system required more development effort than was expected. The reflector type of antenna system finally decided upon is described by Cruser.[1] The antenna directivity at two frequencies is shown in Fig. 12 for both horizontal and vertical polarization. The directivity as well as other characteristics, of course, vary with frequency. The gain of this antenna varies from 6 decibels in the low band to 8 to 10 decibels in the high band. Minor lobes are down about 15 to 20 decibels over most of the band. The coupling between receiving and transmitting antennas on the same mast is a maximum of 26 decibels and decreases to more than 60 decibels when cross-polarization of the antennas is employed.

References

1. Equipment and Mechanical Features of the AN/TRC-24 Radio Set, V. I. Cruser. *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 544–47.

2. Considerations for a New Military Radio Relay System, M. L. Ribe, S. P. Brown. *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 547–52.

◆

No Discussion

Telegraph Terminal AN/FGC-29 Circuit Design Aspects

J. E. BOUGHTWOOD
ASSOCIATE MEMBER AIEE

Synopsis: This paper discusses some of the unique design features of the terminal required to achieve optimum performance on long-range, high-frequency radio circuits. These features are primarily concerned with the receiving terminal and include diversity combining equipment, subgrouping equipment, and delay measuring and equalizing equipment.

TELEGRAPH terminal *AN/FGC-29* was designed to meet the specific requirements of military radio communication. These requirements are necessarily of an exacting nature and the resulting circuitry involves features not commonly employed in commercial telegraph terminal equipment. The importance of securing optimum performance and continuity of operation was a major design consideration. It is the purpose of this paper to cover briefly the more significant characteristics of the equipment circuitry.

The terminal provides 16 frequency-shift 100-word-per-minute telegraph channels starting at 425 cycles and continuing at 170-cycle intervals to 2,975 cycles. Multiplexing equipment is included to permit simultaneous utilization of both 6-kc side bands of the radio facility for telegraph, telephone, or facsimile. The telegraph channel transmitters are conventional in design, employing frequency-shift oscillators such as are commonly used in wire-line systems. The oscillator tank constants are varied by switching diodes so as to produce marking frequency when the sending telegraph loop is closed and spacing frequency when the loop is open. The receiving terminal is considerably more complex than wire-line equipment because of the special requirements imposed by the vagaries of radio propagation.

Diversity Operation

It is common practice on long-range radio-telegraph circuits to employ diversity operation, either space or frequency or both, so as to mitigate the effects of selective fading. A number of schemes have been proposed and utilized by various workers in the field, all for the purpose of extracting the maximum useful intelligence from the received signals.

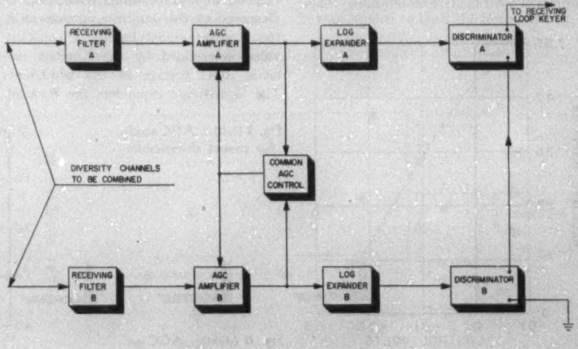


Fig. 1. Ratio squaring combiner

Paper 54-403, recommended by the AIEE Radio Communications Systems and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 18, 1954; made available for printing August 30, 1954.

J. E. Boughtwood is with the Western Union Telegraph Company, New York, N. Y.

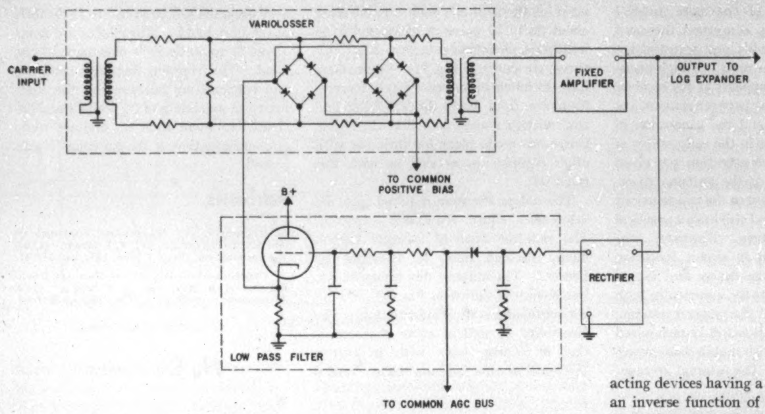


Fig. 2. AGC amplifier

A comprehensive study of a number of combining systems was made for the Signal Corps by Crosby Laboratories, Inc. and resulted in the "ratio squaring" method employed in the present terminal. Briefly the method involves the following basic concepts.

First, the signals to be combined are added subsequent to detection so that the signal voltages add arithmetically in phase while the random noise components associated with each signal can add only on a rms basis. A 3-decibel (db) improvement in signal to noise is thus obtained when two signals of equal amplitude are combined. This is in addition to the normal diversity improvement factor.

Second, to obtain the optimum signal-to-noise ratio when the received signals differ in amplitude, the theory indicates that the signals shall contribute to the combined resultant in proportion to the square of their relative amplitudes. This is of particular importance in the presence of a high noise level where the noise associated with a faded signal would otherwise be combined with the normal signal to effect a 3-db degradation. Further, any amplitude modulation effects on the received carrier should be eliminated, a function normally performed by a limiter in conventional circuits, so as to preserve the advantages of frequency-shift transmission.

Combining Circuit

A circuit operating in accordance with the combining theory is shown diagrammatically in Fig. 1. Common automatic gain control (AGC) amplifiers and individual logarithmic expanders are utilized. The amplifiers, having a common AGC, always have equal gain, the gain being an inverse function of the signal possessing the higher received level. Thus the ratio of the two incoming signal levels is preserved at the amplifier outputs with the stronger signal having a constant value determined by the common delayed AGC voltage of the amplifiers. The logarithmic expanders are forward acting devices having a loss in db which is an inverse function of the input level in db. The amplitude ratio of the signals at the expander outputs thus follows the square of the input ratio as required by the theory. The signals are combined by adding the discriminator outputs in series aiding as indicated.

The design objectives for the combining system contemplated a 45-db fading range over which the common AGC amplifiers should maintain the same relative gain to close limits so as to fully realize the theoretical advantage of the ratio squaring method. Additionally the response time was required to be sufficiently short to remove the amplitude effects of noise as well as the amplitude modulation introduced by channel filters in the presence of keyed signals. The bandwidth of the channel filters, particularly the receiving filter, is the major factor determining the time constant of the AGC system. Filters having a high discrimination against adjacent channels are inherently narrower than less rigorous filters and so introduce a greater amplitude-modulation transient on frequency-shift signals. The amplitude transient is in the form of a damped oscillation having a frequency of approximately twice the maximum dotting rate, or about 80 cycles

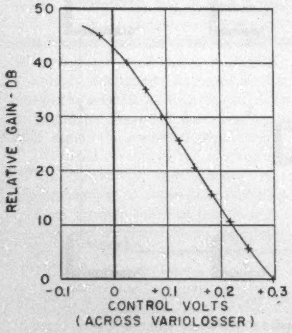


Fig. 3 (left). AGC amplifier control characteristic

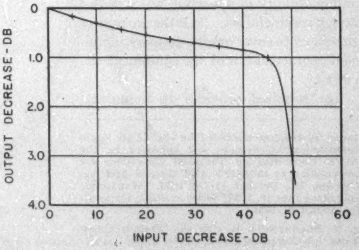


Fig. 4 (right). AGC amplifier level response

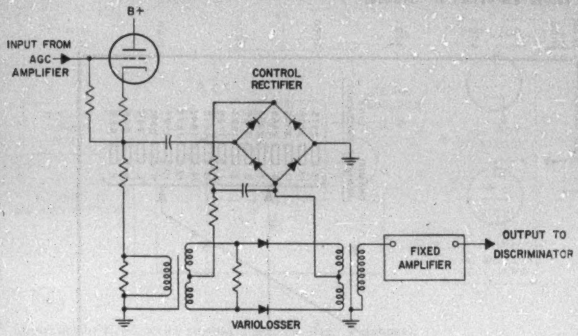


Fig. 5. Logarithmic expander

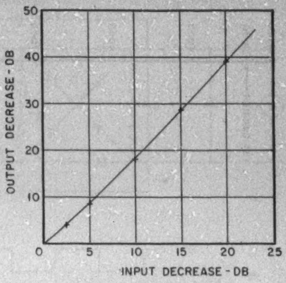


Fig. 6. Logarithmic expander level response

in this case. This is also about the maximum cancelation rate between signal and noise components. In addition, the AGC system must not only compensate for signal fades of 45 db but also must not interfere with the operation of the lowest frequency telegraph channel at 425 cycles, and must have an adequate margin of stability against self-oscillation under all conditions.

It may be of interest to indicate some of the conditions which must be met if a stable system is to be realized. Three general cases are given. An AGC amplifier can be considered as a variable attenuator or variolosser followed by a fixed gain stage, a detector, and a low-pass filter feeding the d-c control voltage back to the variolosser as indicated in Fig. 2. The variolosser can also be considered as a balanced modulator wherein the amplitude of the incoming signal is modulated in accordance with variations in the d-c control voltage.

Case 1—Low-Frequency Stability

Low-frequency stability involves the direct regeneration of control frequencies lying below the cutoff frequency of the low-pass filter. Oscillation will occur where the loop gain is greater than unity for any frequency having a phase shift of 180 degrees. Loop gain is the sum of the fixed amplifier gain, the low-pass filter loss and the magnitude and phase of the modulator unbalance. Theoretically a perfectly balanced modulator would provide infinite loop loss but practically, to maintain a high degree of balance over a fading range of 45 db, would require extremely close matching of components. This is undesirable from an operational viewpoint where the ability to replace components from stock items is a practical requirement. It is therefore necessary to control the loop characteristics by proper design of the low-frequency cutoff of the fixed amplifier and the high-frequency cutoff of the low-pass filter to meet the attenuation and phase-shift characteristics for stable operation.

Case 2—Subharmonic Generation

Instability of this type involves the received carrier frequency, the action being similar to a 2-to-1 regenerative frequency divider. If F represents the received carrier frequency and the loop loss for $F/2$ is insufficient, then the $F/2$ component in the output of the low-pass filter will regenerate itself in the modulator stage by interaction with F, and both F and $F/2$ will appear in the amplifier output.

Case 3—Modulation Instability

Instability of this type produces an amplitude modulation on the received carrier. Modulation takes place in the balanced modulator or variolosser stage and is subject to the full gain of the fixed stage. The modulation is recovered in the detector and, if the low-pass filter attenuation is insufficient and its phase shift is 180 degrees, the modulation will be self-sustaining. As the phase shift is decreased, a damped modulation transient will be produced each time the carrier shifts due to signalling. The decrement of the transient increases as the phase shift is further reduced.

The requirements for stability and the rapid response time needed to simulate limiter action indicated that the desired performance was incompatible with a carrier frequency as low as 425 cycles. It was decided to limit the lowest carrier

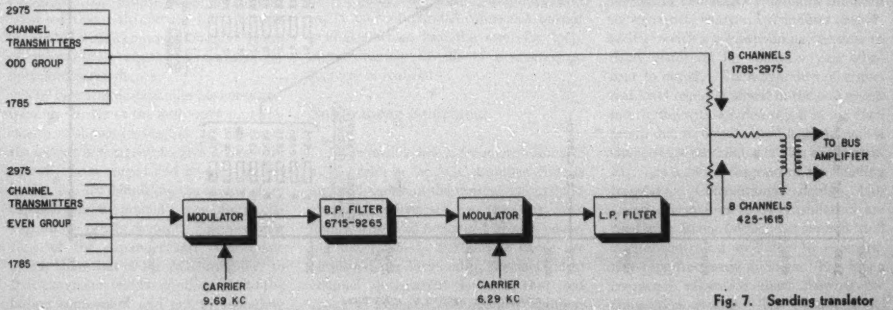


Fig. 7. Sending translator

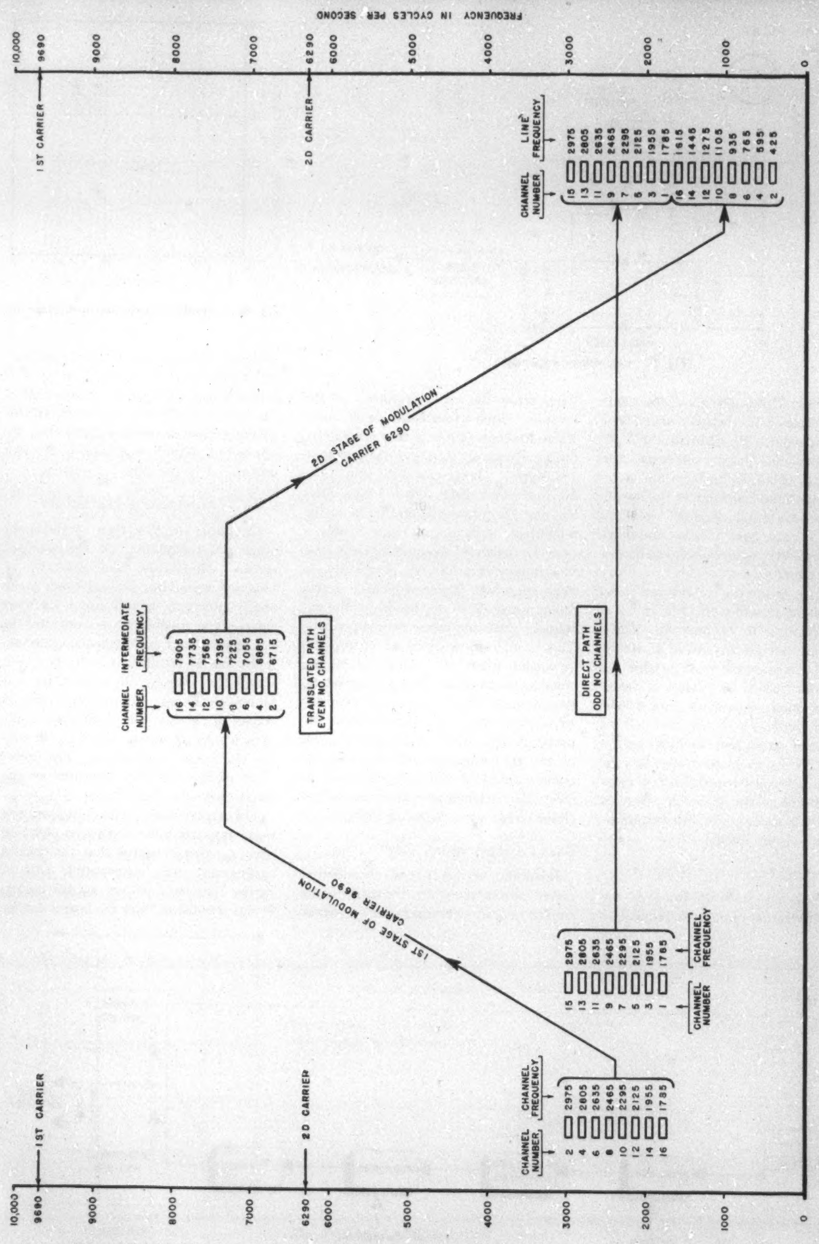

FREQUENCY IN CYCLES PER SECOND
10,000
9690
9000
8000
7000
6290
6000
5000
4000
3000
2000
1000
0
1ST CARRIER
2D CARRIER
CHANNEL NUMBER
CHANNEL FREQUENCY
2 4 6 8 10 12 14 16
2975 2805 2635 2465 2295 2125 1955 1785
15 13 11 9 7 5 3 1
1ST STAGE OF MODULATION CARRIER 9690
CHANNEL NUMBER
INTERMEDIATE FREQUENCY
16 14 12 10 8 6 4 2
7905 7735 7565 7395 7225 7055 6885 6715
TRANSLATED PATH EVEN NO. CHANNELS
2D STAGE OF MODULATION CARRIER 6290
DIRECT PATH ODD NO. CHANNELS
LINE FREQUENCY
15 13 11 9 7 5 3 1 16 14 12 10 8 6 4 2
2975 2805 2635 2465 2295 2125 1955 1785 1615 1445 1275 1105 935 765 595 425

Fig. 8. Sending translator frequency allocations

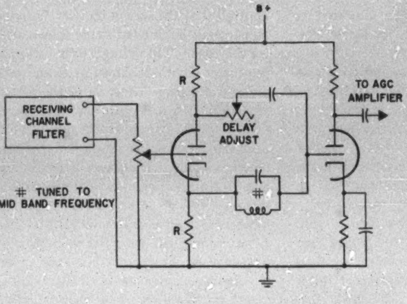


Fig. 9 (above). Adjustable delay network

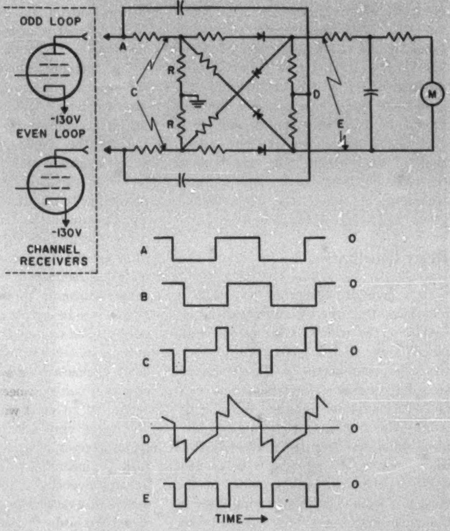


Fig. 10 (right). Differential delay indicator

frequency to 1,785 cycles so that a satisfactory network design could be achieved. As shown in Fig. 2, the variable gain section of the AGC amplifier is a 2-stage balanced variolosser employing germanium diodes as variable impedance elements. The variolosser is designed to have a linear loss in db relative to applied control voltage so that substantially uniform AGC sensitivity is realized over the fading range with corresponding uniformity of transient response. Fig. 3 shows the relation between amplifier gain and applied control voltage. Fig. 4 shows the relation between the input and output levels of the AGC amplifier. A 45-db reduction in input produces only a 1-db reduction in output (a voltage ratio of 1,600 to 1), so that the equivalent of limiter performance is substantially realized.

It is possible to maintain a much closer tolerance on the gain characteristic with varistor variolossers than with conventional AGC circuits utilizing vacuum tubes. This is particularly important for 4-channel combining where the four AGC amplifier gains must track over the fading range if optimum circuit performance is to be realized. If the maximum operating temperature is excessive for germanium, silicon junction diodes can be employed. In either case precautions must be taken to operate the varistors associated with a common diversity group at substantially equal temperatures for optimum performance.

The logarithmic expander also uses germanium diodes as the loss controlling elements as shown in Fig. 5. In this case, the circuit is designed to give a 2-to-1 db ratio between output and input as shown in Fig. 6. By definition, the action of an expander is to magnify any amplitude variations of its input. The response time of the expander has been made slower than that of the AGC amplifier so that expansion will be applied only to the fading component and not to signalling transients or high-frequency cancellations between signal and noise. Here again silicon junction diodes can be used if temperature conditions warrant with a slight change in circuit configuration. Advantage can be taken of the fact that the loss characteristic of a silicon variolosser is less influenced by temperature than germanium when the control current is essentially independent of diode resistance. This could not be done with the AGC amplifier where the variolosser control current had to be an inverse function of diode resistance in order to obtain a linear instead of logarithmic response.

Addition of the diversity signals is effected subsequent to detection by connecting the discriminator outputs in series aiding. The circuitry is conventional and includes a d-c limiting amplifier driving a pair of parallel-connected output pentodes which key the receiving loop. Approximately 25 db of postdetection limiting is realized.

Subgrouping Equipment

To permit a lower frequency limit of 1,785 cycles in the AGC amplifier circuits and still maintain the normal 425 to 2,975 channel frequencies on the line, two identical sets of 8-channel groups covering the 1,785- to 2,975-cycle range are employed per terminal. The method of channel designation is such that one group is assigned odd channel numbers and the other group even channel numbers. The odd group is transmitted and received without translation. The second or even channel group is translated down to the 425 to 1,615-cycle range for transmission. A second translation at the receiver restores the even channels to the 1,785- to 2,975-cycle range. Not only does this method simplify the AGC amplifier design but it also reduces the fortuitous keying loss usually associated with low carrier frequency channels. A secondary advantage is the reduction in number of channel filter types and other tuned circuits peculiar to a given frequency channel.

The group translators employ two successive stages of modulation to achieve the desired frequency conversion. Conversion in two stages eliminates the need for extremely sharp cutoff in filter characteristics with a corresponding decrease in delay distortion for channels lying adjacent to cutoff. The possibility of a normal level received signal in the odd group interfering with a faded signal in the even group due to direct signal leak through a single-stage modulator is also eliminated. Fig. 7 is a block diagram of the sending translator. Conventional double balanced germanium diode modulators are used with carrier frequencies derived from crystal-controlled oscillators by regenerative-type frequency dividers. Fig. 8 is a frequency allocation chart showing the translation process. The even channel group is first modulated by a carrier of 9.69 kc. A band-pass filter selects the lower side band lying between 6,715 and 7,905 cycles. This side band is then modulated by a 6.29-kc carrier. The lower side band selected by a low-pass filter now lies between 425 and 1,615 cycles in the normal line frequency allocations of the system. The translator at the receiving terminal employs a similar technique to restore the even channel group to its original frequency band.

Delay Equalization

In a diversity telegraph system, it is important that the discriminator output signals to be added be in proper time phase. Any time displacement of diversity channels inherent in the system design due to filters, interconnecting wire facilities, or radio equipment will directly reduce the tolerance to fortuitous displacements resulting from selective fading or noise. An adjustable delay network has therefore been provided at the output of each channel receiving filter so that systematic delay differences between channels to be combined may be equalized. A maximum delay of 6 to 7 milliseconds is realized by the bridge-type phase shifter shown in Fig. 9. This amount of delay is easily achieved because of the narrow bandwidth allocated to each channel relative to its carrier frequency.

The measurement of systematic delay differences introduced by transmitting and receiving equipment separated by an intervening radio circuit is complicated by the effects of selective fading which causes random time shifts in the receiving signals and may cause them momentarily to fall below the noise level. To alleviate this difficulty the differential delay indicator shown in Fig. 10 has been incorporated in the receiving terminal metering facilities. In practice a single dotter transmits a-c reversals on the two channels to be equalized in respect to each other. The two channels are terminated in separate loops which are patched to the delay comparator. The comparator consists of a pair of balanced bridge arms *R* connected to an integrating zero center d-c meter circuit through a switching-type detector operated in synchronism with the incoming loop reversals. If the dotting reversals on the two loops are received in exact time phase, the voltage across the bridge arms *R* will be zero and the meter will read zero. When the two loop signals have the time relationship indicated at *A* and *B*, the signal appearing across the bridge will be as shown at *C*. The diode reversing switch is in effect controlled by the sum of the differentiated loop signals having the shape shown at *D* for this case. The voltage to the integrating meter circuit is shown at *E* and produces a negative deflection of the meter. If the time relation of the two loops is reversed, the bridge and switch output voltages also reverse to produce a positive meter reading. Random jitter of the loop transitions due to fading or noise will integrate to zero so that only systematic differences in delay will be indicated. The signal having the least delay is brought into proper alignment by means of its adjustable delay network. Delay adjustments are required to be checked when changes in the system arrangement are made and not as a routine operation.

Conclusions

Telegraph terminal *AN/FGC-29* is not yet in production but comprehensive tests have been conducted on a design approval model to evaluate the electrical design and substantiate the combining theory. A multiple path fading simulator and noise generator were employed to simulate a long-range radio circuit. Performance was definitely superior to currently available equipment but qualitative data must await further testing on actual circuits.

No Discussion

Telegraph Terminal AN/FGC-29 Equipment Features

F. H. CUSACK
MEMBER AIEE

Synopsis: This paper discusses briefly the technical characteristics of telegraph terminal *AN/FGC-29* and the technical requirements that had to be met in its design. The type of construction employed in the equipment is described with the aid of illustrations of a design approval model.

TELEGRAPH terminal *AN/FGC-29* is the terminal equipment for a voice-frequency carrier telegraph system, designed in accordance with military requirements and intended for use in transmission over long-distance military radio circuits. The equipment is being developed by Western Union under a Signal Corps contract at the present time. So far, the developmental phase of the project has been completed and a design approval model constructed. This was accepted after extensive tests by Signal Corps engineers, and a final model now under construction will incorporate certain technical changes found necessary in the course of that work. Up to the present time, no operating experience has been obtained with the *FGC-29* terminal in field service, so the information on the subject that can be presented now is in the nature of a preliminary report on the project. However, the results of the acceptance tests indicate that the equipment will meet all of the specified requirements and that the system will provide dependable operation under extremely adverse conditions of selective radio fading.

Technical Characteristics

Telegraph terminal *AN/FGC-29* provides facilities for 16 carrier telegraph channels capable of operation at speeds up to 100 words per minute. The telegraph channels use frequency shift carrier signals with a deviation of ±42.5 cycles from the mean channel frequency. The center frequencies of the channels are spaced 170 cycles apart, starting at 425 cycles per second and continuing up to 2,975 cycles per second.

Electronic keying is employed on both the sending and the receiving telegraph loops. The loops use neutral d-c signal

---

Paper 54-402, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 18, 1954; made available for printing September 2, 1954.

F. H. Cusack is with the Western Union Telegraph Company, New York, N. Y.

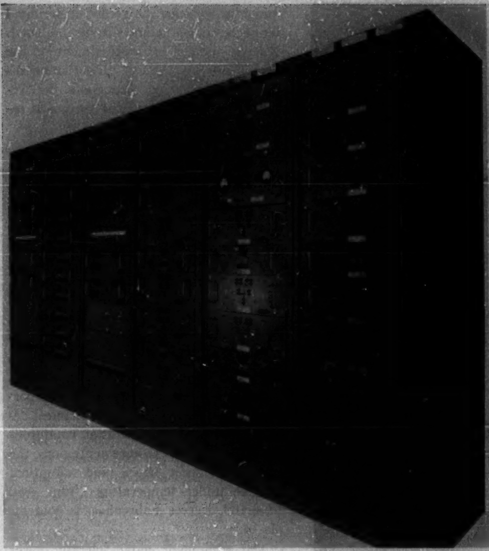


Fig. 1. Telegraph terminal AN/FGC-29, design approval model

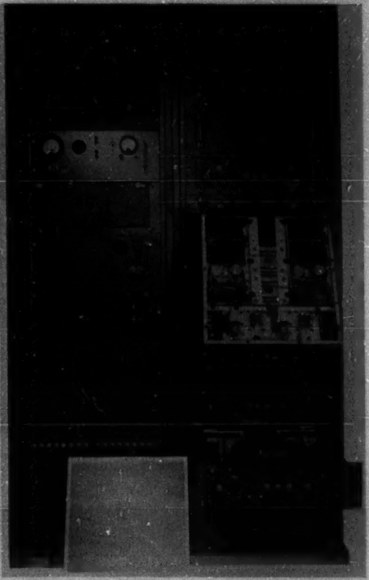


Fig. 3. Transmitting terminal showing drawer-type chassis

current of either 20 milliamperes or 60 milliamperes. Loop power supplies are provided in the equipment, and the transmitting terminal will also meet situations where loop battery of either polarity is to be furnished by the associated telegraph equipment connected to the loop.

The *FGC-29* terminal will be used to provide teletype communication over twin-channel, high-frequency, single-sideband radio circuits. Operation through selective fading is accomplished by means of diversity combining methods with facilities provided for both 2-channel and 4-channel diversity combining, using frequency diversity, or space diversity, or both.

The performance requirements for 2-channel diversity combining are that the over-all peak telegraph distortion introduced by the terminal equipment shall not exceed 5 per cent at 100 words per minute, provided either of the incoming diversity signals is being received at normal level. If one signal fails completely and the second is attenuated by as much as 40 decibels, the peak distortion shall not exceed 7 per cent. These requirements have been met despite the unusual degree of discrimination against interchannel interference which had to be provided in the receiving channel filters.

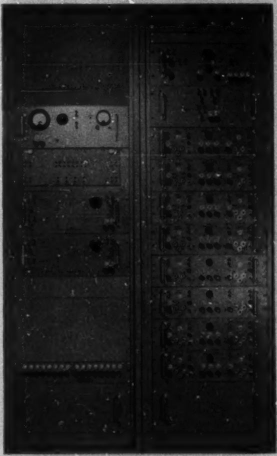


Fig. 2. Transmitting terminal

Because of the selective nature of radio fading, it is required that any channel be capable of satisfactory operation even though the fading produces a 40-decibel difference in level between channels.

In addition to carrier telegraph facilities, the complete terminal includes multiplexing equipment to derive two 3-kc voice-frequency circuits from each of the 6-kc side bands of the radio. Equalizing and amplifying equipment is also provided to permit operation of the terminal over cable pairs to remotely located radio stations.

Diversity Combining

The *FGC-29* equipment employs a unique method of diversity combining which provides a high degree of system flexibility. The incoming signals which are received over two different space diversity or frequency diversity paths are applied to two separate inputs of the receiving terminal, designated as the *A* path and the *B* path. For each telegraph channel, signals from the *A* path are combined with those from the *B* path to provide a single output. With 2-channel diversity combining, the signals so combined are always of the same frequency.

Under adverse fading conditions, further operational improvement is obtainable by employing 4-channel diversity combining. This is done in accordance with a prearranged system whereby

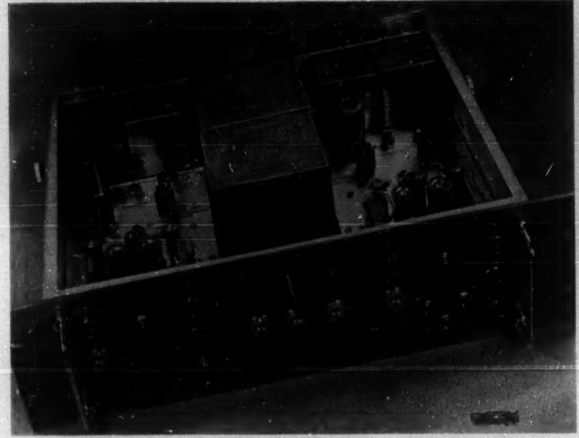


Fig. 4. Transmitter drawer

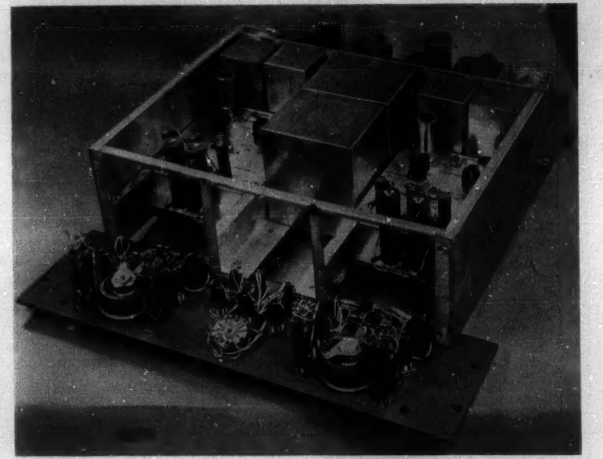


Fig. 5. Transmitter drawer, front panel dropped

the even-numbered telegraph loops are made idle and each of the odd-numbered transmitting loops is made to drive two channel transmitters simutaneously. For example, if channels 1 and 2 are switched to 4-channel diversity, transmitting loop 2 becomes idle and the telegraph signals on loop 1 drive both channel transmitters together. At the receiving end, four incoming signals are combined to produce one output. Two of these are at the frequency of channel 1, and the other two at the frequency of channel 2. The output telegraph signals appear on receiving loop 1, while loop 2 becomes idle. The channel frequency allocation is such that the odd-numbered and even-numbered channels to be combined are always 1,360 cycles apart in frequency. Transferring from 2-channel to 4-channel diversity combining is accomplished by operating one switch in the transmitting terminal and one in the receiving terminal for each pair of channels to be combined. No realignment or readjustment of the equipment is necessary. If 4-channel diversity is employed throughout the terminal, the capacity of the system is reduced to eight channels.

Equipment Description

The *AN/FGC-29* terminal is designed specifically for military fixed plant installations where it must provide long periods of reliable service with a minimum amount of maintenance and readjustment. Service conditions include ambient temperatures ranging from 32 to 122 degrees Fahrenheit and relative humidities up to 95 per cent. Ruggedness and the ability to withstand the rigorous handling encountered in military service are important considerations in the mechanical design. Military transportation conditions impose unusual requirements on the ability to resist shock and vibration. Storage under temperatures from −80 to +160 degrees Fahrenheit is contemplated.

The equipment must tolerate power supply variations of ±10 per cent from the nominal 230 or 115 volts a-c. Regulated power is provided in the terminal wherever needed to meet this requirement. There are three alternating voltage regulators, each with a capacity of 500 watts. Total power consumption for the terminal is 3,500 watts.

As shown in Fig. 1, the complete terminal consists of six steel cabinets, each 75 inches high, $22^1/_2$ inches wide, and 24 inches deep. Each cabinet is ventilated by a blower with filtered air intake. The first two cabinets contain the transmitting equipment and the remaining four comprise the receiving equipment. The sending and receiving equipments are completely independent of each other and need not be installed in the same location. The installation work required for the terminal has been made very simple and is largely a matter of running in the a-c power and the telegraph loop circuits. All interconnections between cabinets and between units within a cabinet are quickly made by means of plug-in multiconductor cables which permit installation without resorting to wiring lists and without making soldered connections.

Fig. 2 shows the transmitting terminal alone. Carrier transmitters for the 16 telegraph channels are contained in eight panels mounted in the cabinet on the right. Each panel contains the two transmitters which are associated with each other for 4-channel diversity operation.

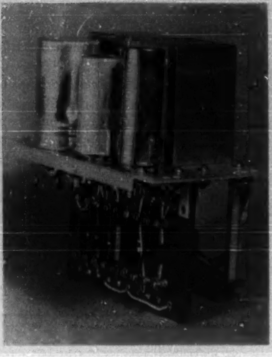


Fig. 6. Typical subassembly

Metering and jacking facilities required for the operation of the terminal are contained in the cabinet on the left, along with multiplexing and power supply equipment.

As shown in Fig. 3, drawer-type chassis construction is employed so as to afford the maximum accessibility for maintenance. Drawers may be drawn out to permit work on the top or bottom of a chassis without interrupting service. Those drawers which are associated only with specific channels may be removed from the cabinet without interfering with the rest of the system. The drawers containing multiplexing and equalizer amplifier equipment are independent units, complete with power supply. When needed at a remote location, they are removed from the drawer slides in the cabinet and are then mounted by their front panels on any standard 19-inch relay rack. No wiring changes are required in the terminal when this is done. The missing units are replaced by dummy circuit panels provided with connectors which cut through the transmission circuits from which the equipment was removed.

Fig. 4 shows one of the channel transmitter drawers and will serve as a representative example of the chassis construction employed throughout the *FGC-29* equipment. This particular drawer contains the carrier telegraph transmitters for channel 1 and channel 2. When the diversity switch in the center of the front panel is in the 2-channel position, the two transmitters are keyed independently by d-c telegraph signals from loop 1 and loop 2. But when this switch is thrown to the 4-channel position, both transmitters are keyed from loop 1 in the manner explained previously. Also located on the front panel are carrier and loop jacks, fuses and blown fuse indicators, and such other controls as are needed for day-to-day operation and maintenance practices. Test points and controls that are used less frequently are located inside the drawer and become accessible when the drawer is pulled out.

Fig. 5 shows the same drawer with the front panel dropped down to reveal the internal construction. It will be seen that the drawer chassis forms a framework on which unitized subassemblies are mounted. Filters and similar sealed units are located in the center with electronic subassemblies, such as amplifiers, oscillators, modulators, and the like on either side. Power supplies are mounted in a separate compartment across the rear. External power is brought to a connector at the rear right and external signal wiring to a multiple pin connector at the rear left. Cabling in the bottom of the drawer interconnects these with the various subassemblies and with the components which are located on the front panel.

A typical subassembly is shown in Fig. 6. This is a miniature amplifier used repeatedly throughout the terminal for a variety of different circuit applications. It has a gain of 50 decibels and is stabilized with about 25 decibels of negative feedback. It is used in 600-ohm circuits over the frequency range from 350 to 6,000 cycles per second. Rated power output is 22 decibels above 1 milliwatt, at which level the second and third harmonic distortions are each down by at least 60 decibels.

Three cabinets of the receiving terminal are shown in Fig. 7. Channel receivers for 16 carrier telegraph channels are contained in 16 separate panels, each of which provides for incoming signals from *A* path and *B* path of one channel. In this illustration the receiver for channel 1 is the large panel near the top of the first cabinet. Directly below it is a smaller panel, the combiner for channels 1 and 2, and below that is the receiver for channel 2. These three units perform the receiving function corresponding to that of the transmitter drawer shown in

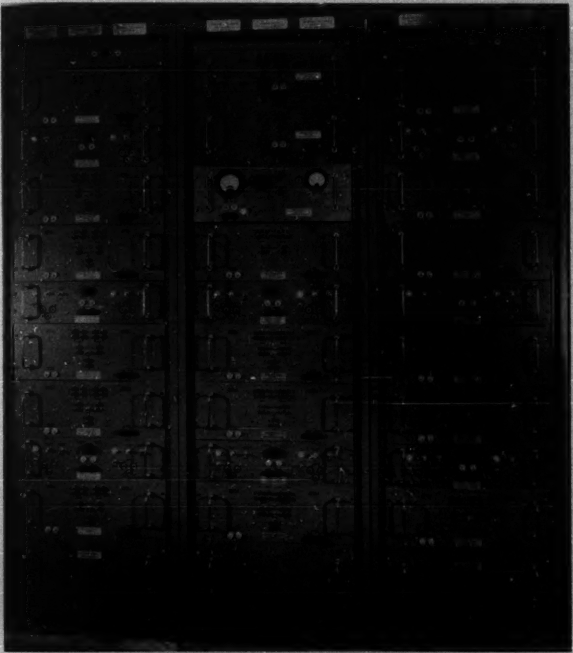


Fig. 7. Part of receiving terminal

Fig. 4. When the diversity switch on the front panel of the combiner is in the 2-channel position, incoming signals from the *A* path and *B* path of channel 1 are combined to provide d-c telegraph signals for loop 1. The process is repeated independently for channel 2 and loop 2. But when this switch is in the 4-channel position, all four inputs are combined to key loop 1 only. The arrangement described for channels 1 and 2 is repeated for all the rest of the channels in pairs, each pair consisting of an odd-numbered channel and an even-numbered channel.

The receiving terminal has its own metering facilities, including a meter for measuring the difference in propagation time for the signals which are combined. Adjustable time-delay networks in the receivers permit these differences to be equalized. Also included in the receiver are jacking and alignment facilities, multiplexing and equalizer amplifier equipment, and power supplies contained in the fourth cabinet, which is not shown in this illustration.

Conclusion

This brief discussion is no more than a preview of telegraph terminal *AN/FGC-29*. Only the more significant features of the equipment have been described and its more important characteristics mentioned. Because this terminal is quite versatile in its applications and can be adapted to a number of different situations, it is expected that the *FGC-29* will be used widely in establishing future military communication networks.

No Discussion

A New Multichannel Teletype Terminal for Use on Long-Range High-Frequency Radio Systems

ALFRED MACK
NONMEMBER AIEE

R. H. LEVINE
ASSOCIATE MEMBER AIEE

THIS paper will describe, on a systems basis, a new multichannel teletype terminal which has been designed for the main-line intercontinental radio circuits of the Army Command and Administrative Network (ACAN). This network provides the backbone of the traffic-handling facilities between the United States and major overseas commands largely utilizing single-side-band transmission and reception. The continuity of communications required in this service makes reliability of paramount importance. Therefore, the equipment must be capable of operating effectively under adverse conditions of long-range radio propagation. The heavy traffic load necessary in maintaining the extensive overseas forces imposes the additional major requirement of maximum communication channels per radio circuit. To meet the varied requirements for teletypewriter, voice and facsimile channels, and the rigors of service in poorly controlled environments sometimes necessary in military operations, a basically new multichannel terminal was required for this service. A glance at the older type of terminal equipment shown in Fig. 1, and which currently is in standard use, will quickly show the inadequacies which led to this new design.

At the start of World War II, this adaptation of the Western Electric Company type-*40C1* voice-frequency carrier telegraph equipment was in use by the Army to provide up to six channels of teletype over a single radio circuit. This system is still in use and employs 2-tone frequency diversity for each channel mark and space. From Fig. 1, at least one major conclusion is apparent: the ten bays of equipment are quite a space outlay for six 60-words-per-minute (wpm) teletype channels. Also, the equipment is by no means adaptable in regard to temperature and humidity. It is essentially commercial-type, fixed plant gear, and must be treated as such if it is to perform properly.

The six teletypewriter channels are transmitted over one radio side band using 24 audiotones. Two tones are transmitted for each channel, one mark and one space tone. To obtain frequency diversity, double modulation tones are applied which then provide a total of four tones per channel, i.e., the basic voice-frequency band of 12 tones is duplicated in an additional band of 12 tones at somewhat higher frequency. Thus, there are two separate channels of intelligence for each message. The receiving equipment suppresses the contribution of the weaker received channel, either the "normal" or "diversity" tones, before conversion to the d-c teletypewriter signals. For high-frequency radio transmission, diversity operation is imperative for error-free performance.

Because frequency diversity is wasteful of radio frequency spectrum and since the radio transmitter must be operated at a reduced power per channel, a method of space diversity single-side-band operation has been devised. This method, by substitution of space diversity for frequency diversity, eliminates the necessity for transmission of a double band of frequencies and thereby reduces the required bandwidth of the radio-frequency spectrum by 50 per cent. There is also a resultant increase in signal power per channel. This yields an improvement in signal-to-noise ratio at the receiving terminal. Alternately, six additional teletype channels can be placed in that portion of the spectrum formerly occupied by the "double modulation tones" of frequency diversity.

It is notable that, at best, even when modified for space diversity operation, the type-*40C1* equipment occupies an entire 6-kc radio side band with only 12 60-wpm teletype channels. This represents a total capacity of 720 wpm. In practice, the remaining radio side band is split by various types of equipment, comprising another complete terminal assembly, to obtain a low-quality voice order wire and a subcarrier facsimile band.

As indicated in the preceding discussion, all of the currently employed single-side-band radio equipment provides two independent 6-kc side bands. These side bands are fed separately from voice-frequency group modulating equipment ap-

---

Paper 54-401, recommended by the AIEE Radio Communications Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 18, 1954; made available for printing September 2, 1954.

ALFRED MACK AND R. H. LEVINE are with the Signal Corps Engineering Laboratories, Fort Monmouth, N. J.

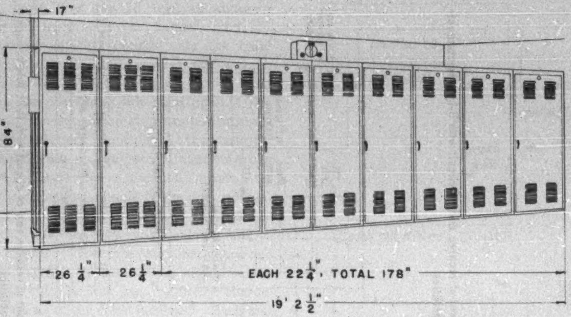


Fig. 1. World War II telegraph terminal for single-side-band high-frequency radio circuit

propriate to the application. The reason for 6-kc channels stems from the requirements of the telephone company's design of the original, and later, single-side-band radio equipment for use in multichannel telephone service. This has also become the standard for all fixed plane type of military single-side-band radio equipments. Therefore, the design of terminal equipment for use over the single-side-band facilities of the ACAN is logically directed towards the maximum utilization of all of the available 6 kc of each radio side band.

While 6-kc signal bands are required for modulation of the radio channels, these same wide bands present a serious problem in audioline facilities. Satisfactory operation of fixed plant radio systems utilizing relatively high-power transmitting equipment requires appreciable separation of transmitter and receiver stations. Also, the communications center, where the traffic originates, terminates, or is tape-relayed, is by its nature usually located in a major headquarters in a large city or on a major military reservation. Therefore, the radio receiving and transmitting stations must be located where space is available for antennas and radio interference is minimized. Telephone line facilities, either metallic or carrier, which are capable of accommodating 6-kc bands are both expensive and difficult to obtain and most radio relay circuits in standard Army use do not provide for 6-kc channels. Hence another design consideration which the new terminal equipment must meet is the ability to provide full service over standard 3-kc wire and radio relay channels. The fundamental requirements which governed the design of the new equipment for the ACAN multichannel single-side-band radio system are:

1. Reliability.
2. Adaptability to service conditions.
3. Increase of traffic capacity.
4. Maximum use of spectrum.
5. Maximum utilization of 6-kc radio side bands.
6. Operation over ordinary line facilities.
7. Frequency or space diversity options.
8. Ease of installation and maintenance.

Item no. 7, while not mentioned previously, is a required variation for stations which, because of limited space or other considerations, cannot provide diversity radio receiving antennas. Item no. 8 is an obvious requirement for any equipment designed for military employment. It might be noted that the type-*40C1* equipment requires weeks of installation time.

The background and requirements are now firmly established for proceeding to the telegraph terminal *AN/FGC-29* which is the new multichannel terminal specifically engineered to meet the required specifications. Briefly, the basic *AN/FGC-29* provides the following:

1. Sixteen 100-wpm teletype channels.
2. Optional space or frequency diversity.
3. Band-splitting equipment to derive 3-kc voice bands.
4. Necessary equalizer amplifier facilities for operation over standard loaded or nonloaded cable or commercial telephone carrier systems.
5. Compact design.
6. Quick installation.
7. A highly effective diversity combining system.
8. Quadruple diversity for use under adverse conditions of radio propagation.

Fig. 2 illustrates that the telegraph terminal *AN/FGC-29* comprises approximately half the number of bays required for the type-*40C1* equipment even though it provides more than twice the amount of traffic-handling capacity. Plug-in cables are used to interconnect the various bays which reduces installation time to a matter of days. Fig. 2 shows the design approval model, and illustrates the essential features of the production version. The two bays on the left comprise the transmitter portion of the terminal including everything from d-c printer loops to send-line amplifiers and multiplexers. Similarly, the four receiving bays are complete with amplifier-equalizers and demultiplexers through to d-c loops. Send and receive portions of the terminal are independent and can be separated, if required. The basic application of the equipment is for 16 100-wpm space diversity teletype channels. Figs. 3(A) and (B) illustrate the equipment arrangement for this mode of operation at the send and receive portions of the circuit respectively.

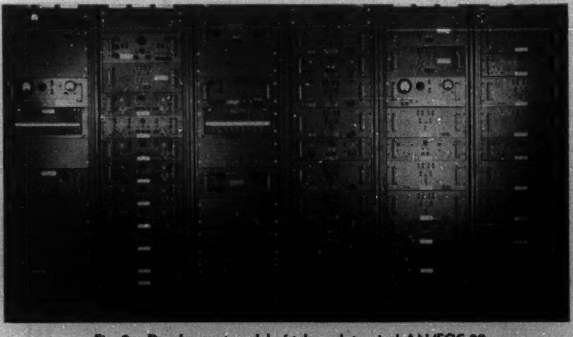


Fig. 2. Development model of telegraph terminal AN/FGC-29

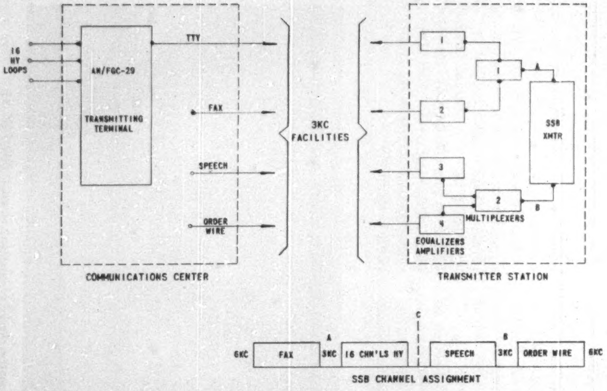

16
HY
LOOPS
AN/FGC-29
TRANSMITTING
TERMINAL
TTY
FAX
SPEECH
ORDER
WIRE
COMMUNICATIONS CENTER
3KC
FACILITIES
1
2
3
4
1
2
A
B
SSB
XMTR
MULTIPLEXERS
EQUALIZERS
AMPLIFIERS
TRANSMITTER STATION
6KC
FAX
A
3KC
16 CHN'LS HY
C
SPEECH
B
3KC
ORDER WIRE
6KC
SSB CHANNEL ASSIGNMENT

Fig. 3(A). Space diversity equipment arrangement at send end of circuit with 3-kc interconnecting facilities

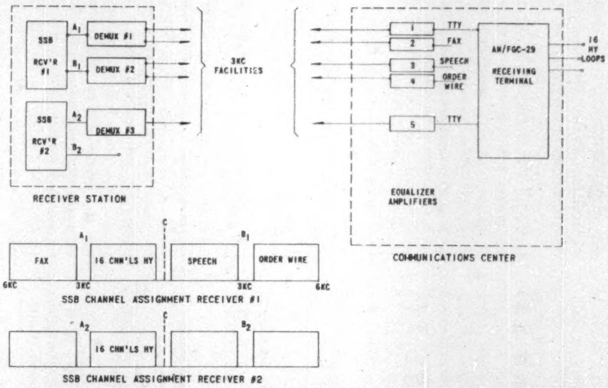

SSB
RCV'R
#1
SSB
RCV'R
#2
A1
B1
A2
B2
DEMUX #1
DEMUX #2
DEMUX #3
RECEIVER STATION
3KC
FACILITIES
1
2
3
4
5
TTY
FAX
SPEECH
ORDER
WIRE
TTY
AN/FGC-29
RECEIVING
TERMINAL
16
HY
LOOPS
EQUALIZER
AMPLIFIERS
COMMUNICATIONS CENTER
A1
C
B1
FAX
16 CHN'LS HY
SPEECH
ORDER WIRE
6KC
3KC
3KC
6KC
SSB CHANNEL ASSIGNMENT RECEIVER #1
A2
C
B2
16 CHN'LS HY
SSB CHANNEL ASSIGNMENT RECEIVER #2

Fig. 3(B). Space diversity equipment arrangement at receive end of circuit with 3-kc interconnecting facilities

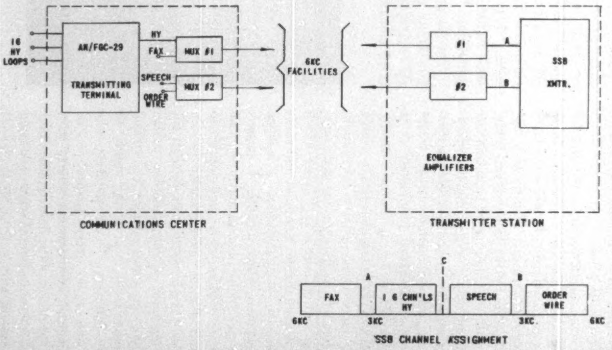

16
HY
LOOPS
AN/FGC-29
TRANSMITTING
TERMINAL
HY
FAX
SPEECH
ORDER
WIRE
MUX #1
MUX #2
COMMUNICATIONS CENTER
6KC
FACILITIES
#1
#2
A
B
SSB
XMTR.
EQUALIZER
AMPLIFIERS
TRANSMITTER STATION
A
C
B
FAX
1 6 CHN'LS
HY
SPEECH
ORDER
WIRE
6KC
3KC
3KC.
6KC
SSB CHANNEL ASSIGNMENT

Fig. 4(A). Space diversity equipment arrangement at send end of circuit with 6-kc interconnecting facilities

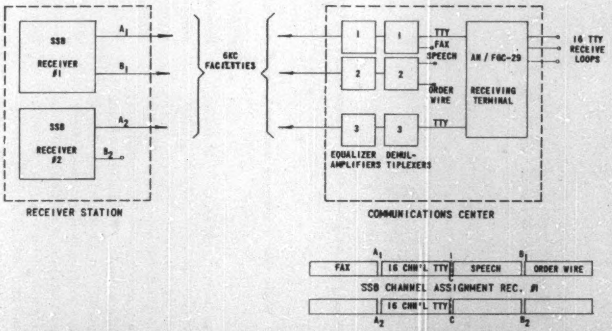

SSB
RECEIVER
#1
SSB
RECEIVER
#2
A1
B1
A2
B2
RECEIVER STATION
6KC
FACILITIES
1
2
3
1
2
3
TTY
FAX
SPEECH
ORDER
WIRE
TTY
AN / FGC-29
RECEIVING
TERMINAL
16 TTY
RECEIVE
LOOPS
EQUALIZER
AMPLIFIERS
DEMULTIPLEXERS
COMMUNICATIONS CENTER
A1
C
B1
FAX
16 CHN'L TTY
SPEECH
ORDER WIRE
SSB CHANNEL ASSIGNMENT REC. #1
16 CHN'L TTY
A2
C
B2

Fig. 4(B). Space diversity equipment arrangement at receive end of circuit with 6-kc interconnecting facilities

Note that the 16 teletype channels which have 170-cycle channel-center frequency spacing occupy only the band from 425 to 2,975 cycles per second. This leaves a remaining 3-kc voice band in the 6-kc radio side-band input. The multiplexer, shown at the radio transmitting station, multiplexes the teletype band with a 3-kc voice circuit into one side-band input of the radio transmitter by means of suitable filters and modulators. The other side band is similarly split to yield two voice bands for use, as required. Note the amplifier-equalizers installed at the transmitter station to correct for the loss and frequency characteristics of the cable facilities. Each amplifier-equalizer is capable of correcting for the loss and frequency characteristics of up to 40 miles of loaded cable or 20 miles of nonloaded cable.

At the receiving station, demultiplexers in each side band of receiver no. 1 provide separation of the various bands of information for transmission over 3-kc audio-line facilities. Each band is restored to its original audio-frequencies before being applied to the transmission facility. Receiver no. 2 is equipped with only one demultiplexer, inasmuch as diversity is normally applied only to the teletype signals. Both sets of teletype bands are fed into the terminal through amplifier-equalizers and diversity combination is effected on an individual channel basis within the terminal equipment. The three remaining channels are amplifier-equalized and routed as required. Note that the teletype channels can be transmitted in any one of four slots in the transmitted side bands. Both diversity outputs are always of the same audio-frequency of 425 to 2,975 cycles per second regardless of the employment of the equipment.

Figs. 4(A) and (B) again illustrate space diversity operation except that here 6-kc facilities are used for the links between the communications center and radio sites. The only change is that the multiplexing equipment is now all located at the communications center. For frequency diversity operation, various setups are possible. The simplest, illustrated by Figs. 5(A) and (B), uses one 6-kc side band for the teletype, leaving the other to be utilized as desired. The basic flexibility of the system also permits placing the teletype in both side bands or, in general, in any two of the four 3-kc channels. Note that one 3-kc circuit has been lost.

Under adverse propagation conditions, improved performance can be obtained by using quadruple diversity. This is done by keying 2-channel transmitters from each send loop in addition to the aforementioned methods of obtaining diversity operation. At the receiving end, the diversity pairs of channels 1 and 2, 3 and 4, etc., are combined in a quadruple diversity selector. This mode may be employed when using either space or frequency diversity operation. It is physically accomplished by front panel selector switches for 2-channel or 4-channel diver-

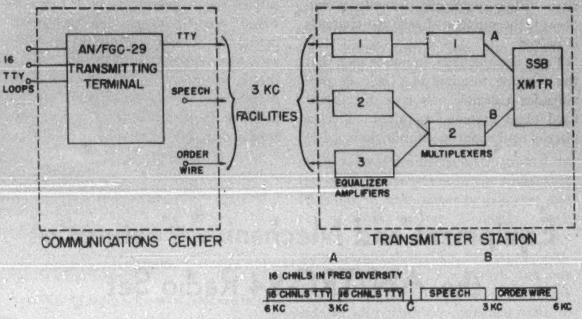

16 TTY LOOPS
AN/FGC-29 TRANSMITTING TERMINAL
TTY
SPEECH
ORDER WIRE
3 KC FACILITIES
1
1
2
2
3
A
B
SSB XMTR
MULTIPLEXERS
EQUALIZER AMPLIFIERS
COMMUNICATIONS CENTER
TRANSMITTER STATION
A
B
16 CHNLS IN FREQ DIVERSITY
16 CHNLS TTY
16 CHNLS TTY
SPEECH
ORDER WIRE
6KC
3KC
C
3KC
6KC

Fig. 5(A). Frequency diversity equipment arrangement at send end of circuit with 3-kc interconnecting facilities

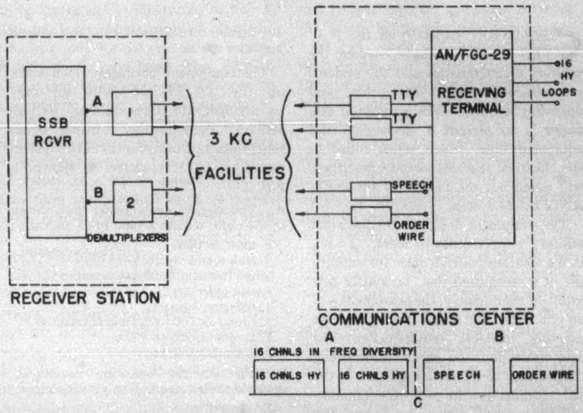

SSB RCVR
A
1
B
2
DEMULTIPLEXERS
RECEIVER STATION
3 KC FACILITIES
TTY
TTY
SPEECH
ORDER WIRE
AN/FGC-29 RECEIVING TERMINAL
16 HY LOOPS
COMMUNICATIONS CENTER
A
B
16 CHNLS IN FREQ DIVERSITY
16 CHNLS HY
16 CHNLS HY
SPEECH
ORDER WIRE
C

Fig. 5(B). Frequency diversity equipment arrangement at receive end of circuit with 3-kc interconnecting facilities

Table I. Comparison of Telegraph Terminal AN/FGC-29 with Existing Terminal Equipment

	40C1			
Item	6-Channel Frequency Division	6-Channel Space Division	12-Channel Space Division	AN/FGC-29
Total wpm	360	360	720	1600
Weight	5220	5220	7465 or more	3000
Size	10 bays 17″ by 84″ by 19′2″	10 bays 17″ by 84″ by 19′2″	15 bays or more 17″ by 84″ by 28′6″	6 bays 24″ by 78″ by 11′6″
Bandwidth	5 kc	3 kc	5 kc	3 kc
Service conditions	commercial fixed plant			military fixed plant
Installation time	several weeks			days
Extras				flexible application, multiplexing equipment, test and meter panels

sity. When quadruple diversity is employed, the number of teletype channels is halved.

From the foregoing, it can be seen that the teletype terminal *AN/FGC-29* provides for a flexible new type of military fixed plant systems operation. It is a piece of equipment specifically designed to satisfy the requirements of the system which existing items have never approached. Tests conducted on the design approval model of the terminal indicate that it will fully live up to expectations. A future paper is planned which will report on the more important performance results obtained. Preliminary results are excellent. A summary of the outstanding features of *AN/FGC-29*, in comparison with the type *40C1* now in use, is shown in Table I.

◆

No Discussion

Equipment and Mechanical Features of the AN/TRC-24 Radio Set

V. I. CRUSER
NONMEMBER AIEE

COMPANION papers[1,2] on the *AN/TRC-24* radio set have covered the military characteristics and the systems requirements, circuit description, and operating features. The purpose of this paper is to present a mechanical and equipment story which, when added to the electrical features already described, will round out the picture of the entire *AN/TRC-24* system.

The equipment designers also have had military requirements to meet. To obtain a well-integrated system for military use, it is recognized that, no matter how excellent the circuitry, the equipment is of no practical value unless it meets certain exacting physical requirements. The equipment must be easily transported, rugged, easy to set up, and quickly on the air. It must operate over wide humidity and temperature ranges day in and day out. It must be easily maintained. And it must be produced at a reasonable cost. In more detail, some of the more rigorous physical requirements are:

1. The packages shall be of such weight and form as to be conveniently handled by not more than two men.
2. Equipment shall withstand 3 hours of bouncing on a table moving vertically at 285 cycles per minute through a total amplitude of 1 inch. This simulates an accelerated jeep ride over rough and frozen ground.
3. Equipment shall withstand storage for an unlimited time in temperatures from −80 to +165 degrees Fahrenheit and operate in temperatures from −65 to +150 degrees Fahrenheit.
4. Equipment shall withstand five 48-hour humidity cycles.

The equipment packages which make up the *AN/TRC-24* radio set may be divided into three groups, as described in the following. A display of actual production units for the entire system consisting of 21 packages is shown in Fig. 1.

A. The antenna and antenna mast, shown at the right, require a total of 11 packages:

3, mast sections } in the foreground
1, stakes and sledge }
1, guy wires and other accessories
1, two reflectors
1, reflector support
1, antennas, 100 to 225 megacycles (mc)
1, antennas, 225 to 400 mc
2, antenna cable reels

B. The essential operating equipment, shown at the left, requires only six packages:

1, transmitter
1, receiver
1, power supply for transmitter
1, alternate tuners
1, autotransformer
1, equipment accessories and spares case

C. Finally there are four supplementary packages:

1, power accessories case including one interconnecting box, one switch box and one radio-frequency wattmeter
1, band-pass filters, 100 to 225 mc
1, band-pass filters, 225 to 400 mc
1, power cable reel

Many of these packages contain mechanical or electronic gear which must be protected from the rigors of the bounce test. To provide this protection, the equipment is housed in an inner case which is supported by means of rubber mounts in an outer or transit case. Units such as the transmitter, the transmitter power supply, and the receiver which contain field replaceable components such as vacuum tubes are arranged so that the complete chassis may be partially withdrawn on roller slides from its inner case for this servicing, as shown in Fig. 2. The construction of these transit cases follows, in general, a design originally developed by the Signal Corps and provides an interlocking feature for stacking.

Both the transmitter and receiver circuits have been sectionalized functionally for convenience of manufacture and field replacement. Each individual subunit is connected to the common wiring by means of a plug and jack to further facilitate shop and field testing. Quick-fastening devices facilitate removing and replacing these units. The tuning units, for both the transmitter and the receiver, are easily replaceable in the field without withdrawing the parent unit from its inner case. Here, particularly sturdy alignment pins control the engagement of the mating plug and jack to guard against damage through careless handling.

Certain details of the antenna design may be of interest. For this design, the simple half-wave dipole and plane reflector was chosen primarily because of its broad-band characteristic and its physical size and shape. Each radiation is a 2-piece telescopic assembly. Only two different designs of dipoles are required to cover the full 100- to 400-mc range. For low-frequency design only three sets of adjustments are needed to cover the 100- to 225-mc range, and for the high-frequency design only two settings cover the entire 225- to 400-mc range. This is important inasmuch as each time the dipoles have to be replaced or readjusted because of a change in operating frequency, the complete mast must be lowered, raised again and reoriented for maximum gain. As illustrated in Fig. 3, both the receiving and sending antennas are mounted atop a single mast. Each reflector is only 4 feet high and 4 feet 6 inches wide. It is hinged in the middle so it may be folded for easy packaging and is arranged to be mounted for either horizontal or vertical polarization of the associated antennas. In the illustration, the antennas at the left are polarized horizontally and those at the right, of lower frequency, are polarized vertically. Stay guys secured to the

---

Paper 54-406, recommended by the AIEE Radio Communication Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11-15, 1954. Manuscript submitted June 18, 1954; made available for printing September 1, 1954.

V. I. Cruser is with the Bell Telephone Laboratories, Inc., Murray Hill, N. J.

Vo
Pt.
19

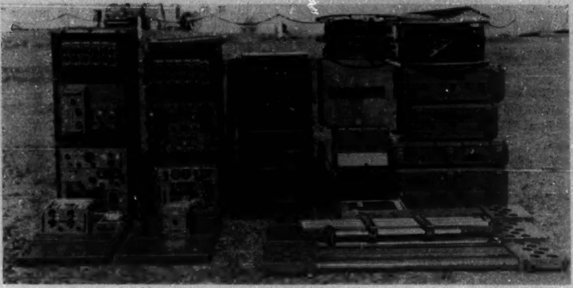


Fig. 1. Complete complement of packages for the AN/TRC-24 radio set, covers removed from transit cases

outside edges of the reflectors stabilize the whole assembly during high winds. The two reflectors with a full complement of dipoles weigh slightly more than 100 pounds.

In some cases separate masts will be required for transmitting and receiving. For these cases, the support bracket may be rotated atop the mast so that only one reflector is centered on each mast. The antenna mast itself is basically of the same design as the one so successfully used all during World War II by the Signal Corps. Normally, nine sections each 5 feet long are fitted together to form a 45-foot mast. A set of four guy wires support the mast at each 15-foot level. The mast and antenna assembly has been designed to withstand a 1/2-inch ice load and 100-mile-per-hour winds.

Of the operating units, the transmitter is most complex from a mechanical viewpoint. For ease of assembly and maintenance, this unit has been designed in three layers and is illustrated in Fig. 4. The top layer, which contains all of the smaller chassis for the generating circuits, is hinged along its lower front edge so that it may be raised to provide access to the middle layer. The middle layer, which contains the blower assembly and the exciter chassis including the driver stage, is stationary. The bottom layer, also stationary, houses mainly the band-pass filter and the final stage, or stages, in the radio-frequency tuning unit. Both of the latter units are arranged to be plugged in from the front.

In the transmitter tuning units, tuned cavities are employed in the final power stages. In the *C*-band unit, these cavities are of the "folded back" or "telescopic" design for compactness so that both the grid and plate cavities are arranged coaxially with the associated high-frequency amplifier tube mounted at the end. Fig. 5 gives some idea of the amount of hardware packed into the 2-stage tuning unit. Note the vacuum tubes held in by the "life-saver" clips at the near end of each stage.

The manner of driving the plate-short-circuiting plungers in the transmitter cavities may be of interest. These plungers, which may be distinguished by the beaded appearance of the contacting tines, are several inches in diameter and travel approximately 5 inches. Each plunger is held in alignment and at the same time is driven by two sturdy outboard lead screws, one above and one below the plunger. These driving screws ride in an accurately threaded yoke to which the plunger is attached. Each pair of lead screws is driven in step from one knob. For this purpose, the simple and rugged sprocket-and-chain drive of the Morse lightweight precision type is employed. With a convenient idler on an eccentric shaft to take up slack, this type of drive contributes essentially no backlash and very little to the torque load.

The design of the contacting tines for the various tuning plungers was a study in itself. The selection of the proper material, the cross section and length of the tine, the shape of the contact, the hardness, the finish, the burnishing, and the lubricant all played a part in getting constant contact with low sliding friction and long life without burning or scoring.

It is in these power cavities that much of the total transmitter input wattage is dissipated. Extreme temperatures at the tube, of course, would not only ruin the tube in no time but would also draw the temper from the associated contact springs. Passageways for cooling air were tortuous in route and very limited in size so that air at a relatively high pressure and volume was required. The

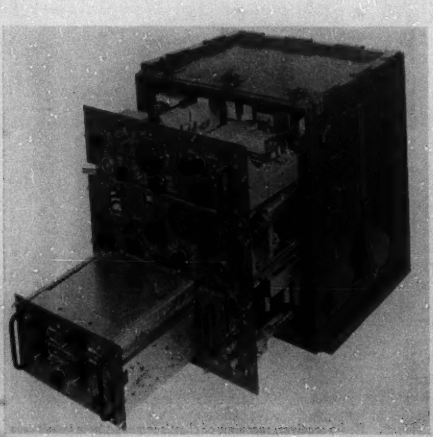
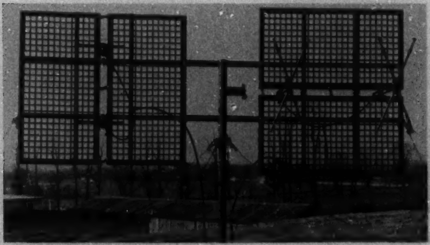


Fig. 2 (left). Radio transmitter, tuning unit and transmitter chassis partly withdrawn from transit case

Fig. 3 (below). Complete antenna shown assembled and mounted on the mast



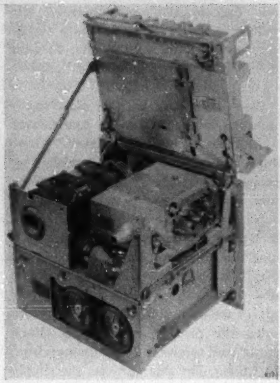


Fig. 4. Radio transmitter, rear view of chassis removed from transit case showing access to blower assembly and exciter unit

package for this air supply had to be small and relatively light in weight. A high-speed a-c vacuum cleaner type of motor with a single stage blower was selected and may be seen in Fig. 4 at the left side of the middle equipment layer. Special attention was given to the shape of the blower fins and housing and to the balancing of the rotating parts to reduce noise and vibration. The motor-blower unit was mounted on rubber.

In spite of the careful balancing and special mounting of the blower, the microphonic pickup in the early stages of the exciter circuit was still intolerable. Resorting to rubber again, the exciter unit, shown at the right of the blower, was mounted on very lightweight vibration mounts. This solved the microphonic problem but, of course, the next step was to immobilize these mounts during the time the equipment was in transit. To accomplish this, an inclined-plane pickup was designed with an operating lever actuated by the front cover of the transit case. This pickup assembly may be seen in Fig. 4 on the under side of the hinged top equipment layer. With the front cover removed from the transit case, the exciter unit rests freely upon its vibration mounts. Replacing the cover actuates the operating lever and the exciter unit is lifted off of the vibration mounts and locked against the framework. The mounts for the blower are made somewhat heavier than normal for vibration damping so that it is unnecessary to lock them out during transit.

Just a few words about the general ventilation of the transmitter. The antigravity system has been used, pulling the cool air in from the top of the transit case and forcing it out the bottom. Utilizing half of the 1-inch space between the inner and outer cases, ducts were formed so that air at relatively low velocity is drawn through a dust filter in the top of the inner case across the vacuum-tube chassis in the top layer of equipment. Then, by means of the blower, it is forced under pressure and at relatively high velocity to the power tubes and then released against atmospheric pressure along the lower edge of the transit case. While, as might be expected, the motor-blower unit develops an appreciable whine, by carefully balancing the rotating parts and using rubber mounts it has been reduced to an acceptable degree. To protect the power tubes in case of blower failure, a centrifugal switch has been built into the motor to cutoff the high-voltage supply when the motor speed falls appreciably.

The receiver was designed so that the separate wired panels are arranged as satellites around the tuning unit as a core. As may be seen in Fig. 6, this arrangement provides easy access for tube replacement by simply withdrawing the whole receiver chassis from its transit case. In fact, all apparatus components recognized as field replaceable items face outward. As in the transmitter, the tuning unit, being optional for either of the two frequency ranges, may be quickly replaced from the front.

The receiver tuning units have been designed around existing types of radio-frequency line tuners. After considerable exploration, a Mallory inductuner was chosen as the heart of each of the receiver tuning units. Fig. 7 shows its use in the 225 to 400-mc receiver tuning unit. Considerable effort went into the development of the "butterfly" type of oscillator circuit associated with this unit. Because of the stringent settability and temperature stability requirements, extreme care was taken to insure that the receiver could be pretuned to the frequency of a remote transmitter. This meant that the butterfly had to be designed with extremely close tolerances and built with exacting care in order to track with a standardized dial. Backlash had to be minimized with close bearing fits and antibacklash gearing; and the assembly had to be of low coefficient ma-

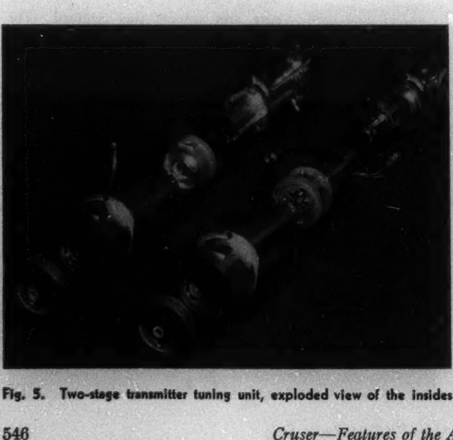


Fig. 5. Two-stage transmitter tuning unit, exploded view of the insides

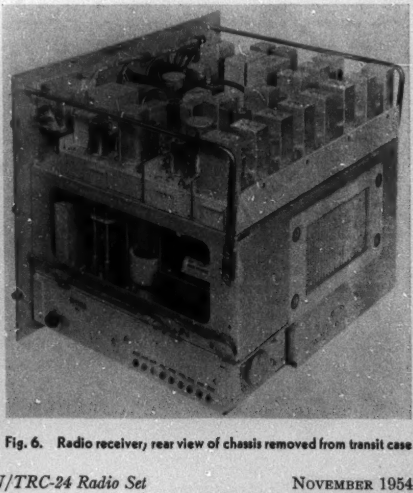


Fig. 6. Radio receiver, rear view of chassis removed from transit case

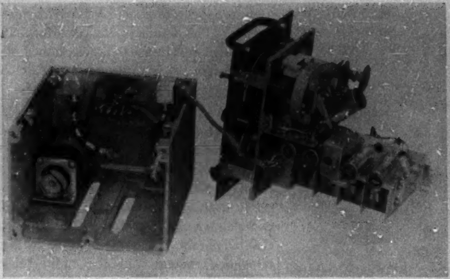


Fig. 7 (above). Receiver tuning unit of 225 to 400 mc showing associated butterfly-type oscillator

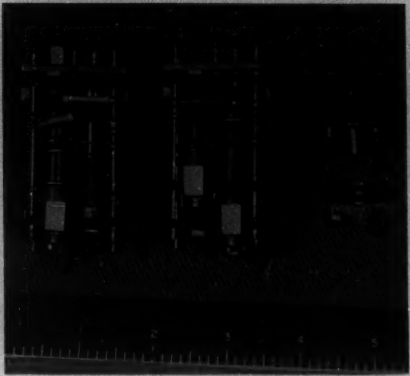


Fig. 8 (right). Two types of network assemblies with adjustable coupling between inductances

terial (invar) with compensation for temperature effects over the full frequency coverage within the range of the automatic frequency control circuit. Note that the oscillator tube is mounted directly upon the butterfly frame to minimize the length of the grid and plate leads. Cooling the receiver involves the use of a small axial-fan type of blower under control of a thermostat to retain self-generated heat when operating at low temperatures. The thermostat starts the fan only after the ambient temperature reaches approximately 85 degrees Fahrenheit.

A series of small components which have played a large part in the design include the intermediate-frequency interstage networks and the limiter-discriminator coils. Most of these involve slug-tuned intercoupled inductances. New designs were called for here since there were no known equivalents on the market which could provide an independent adjustment of the inductive coupling. Two different types of these are illustrated in Fig. 8. It should be noted that the assembly at the left provides end-to-end coils for close coupling; the one at the right provides side-by-side coils for looser coupling. After the coupling is adjusted in the shop, the movable spool heads are soldered to the four corner posts at the eyelets. Final slug adjustments are then made for each coil. While these are simple devices, the man at the test-and-adjust bench appreciates them because they afford him full control of all the factors with which he is dealing.

References

1. A New Ultrahigh-Frequency Multichannel Military Radio Relay System, **John G. Nordahl.** *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 526–31.

2. Considerations for a New Military Radio Relay System, **M. L. Ribe, S. P. Brown.** *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 547–52.

◆

No Discussion

Considerations for a New Military Radio Relay System

M. L. RIBE
NONMEMBER AIEE

S. P. BROWN
NONMEMBER AIEE

Military Requirements for Radio Relay

PRIOR to World War II, the Armed Forces' need for multitelephone channel trunk circuits was satisfied by the use of metallic wire circuits. Early in the war, it was demonstrated that the fluidity of movement which accompanied blitz-type warfare required a more rapid means of providing multichannel trunk facilities than could be had with wire circuits. As a result, military radio relay equipment operating in the very-high-frequency range was developed early in 1942, capable of providing rapidly installable transmission facilities for four frequency-division multiplexed voice channels. This equipment was used successfully in all theaters of operation throughout World War II, conclusively demonstrating the tactical advantages of this type of transmission facility. In addition, 8-channel microwave equipment, using time-division principles of multiplexing, was developed for use at high echelons of command.

Experience gained during World War II demonstrated that fluid movement of large tactical units may again be expected. Furthermore, wide dispersal of Army units was evidenced and could again occur, which will require rapidly installed trunk circuits 100 to 200 miles in length. It may be assumed that situations will exist where tactical units are isolated and the only means for providing multichannel trunk communications will be by radio relay systems. For these

---

Paper 54-404, recommended by the AIEE Radio Communication Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 14, 1954; made available for printing September 14, 1954.

M. L. Ribe and S. P. Brown are with the Signal Corps Engineering Laboratories, Fort Monmouth, N. J.

The efforts of the Western Electric Company's and the Bell Telephone Laboratories' personnel are hereby gratefully acknowledged. The speed and quality of the development and subsequent production efforts are due to the effective teamwork of the engineering and production staffs of these organizations.

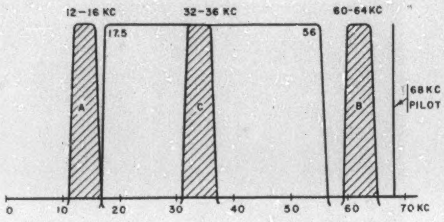


Fig. 1 (left). Interchannel crosstalk test filter characteristics

reasons, radio relay systems will probably provide a primary means for military trunk communications.

Basic System Planning Considerations

The increasing complexity of warfare has resulted in constantly increasing requirements for the quality and capacity of multichannel trunk circuits over those previously available with prior military communications systems. A co-ordinated multichannel system development was therefore initiated by the Signal Corps Engineering Laboratories (SCEL) prior to the Korean conflict for frequency-division multiplexed carrier telephone equipment using either cable or radio relay as transmission mediums. Studies made at the initiation of the program resulted in the decision to initiate development of carrier telephone systems of both 4- and 12-channel capacity. The radio relay system was developed to provide the same quality of transmission as spiral-4 cable systems for either 4- or 12-telephone channel service.

Though several radio relay systems were available at the time the development was started, none was of a design that would meet military requirements. Military use requires that the equipment be capable of working over extreme supply voltage, line frequency, humidity, and temperature ranges, be capable of being moved and installed quickly, stand extended rough usage under tactical field conditions, and be operable by relatively untrained personnel. Furthermore, the need was indicated for wide frequency coverage with a minimum of frequency determining elements to permit the application of flexible frequency allocation plans. Since the area of use and availability of frequency spectra would be changeable, sufficient frequency coverage had to be provided to permit expected heavy multiple system use in a given area with a minimum of interference. The frequency coverage that was required exceeded greatly the tuning ranges of available military or commercial designs. Additionally, the components used to construct the equipment were required to be the standard ones used by the military service, which were already in the military supply systems.

Before the beginning of Korean hostilities, studies conducted by the SCEL had shown that the best design to be used for satisfying existing tactical requirements was for equipment operating in the upper very-high-frequency and lower ultrahigh-frequency ranges, where siting problems were not acute and where fading problems were not expected to be serious. There are many services operating in these frequencies, and it was essential that the radio relay system use the minimum radio-frequency bandwidth consistent with performance requirements. This consideration led to the specification of a system using modified frequency modulation with the pre-emphasis used being a compromise for best crosstalk and noise performance. It was evident that the major design problems were the provision of the wide tuning range required without undue sacrifice of performance, simplicity and reliability, and the maintenance of low amplitude distortion, and therefore low interchannel crosstalk, during field operations.

Experimental work at the SCEL had resulted in a method of generating crystal reference frequencies over wide tuning ranges which promised success, and in circuits which yielded low enough distortion products to meet systems requirements. The details of the circuits which were developed by the Bell Telephone Laboratories for use in radio set *AN/TRC-24* are described in a companion paper.[1]

Based upon existing requirements, the investigations previously mentioned, and experience gained in the use of prior radio relay equipment, specificational requirements were evolved for radio relay equipment usable with either the 4- or 12-channel telephone system. To expedite the availability of both the wire carrier telephone system and the radio relay system, contracts were initiated with the Western Electric Company in July 1950 for both the radio relay and carrier components. Many relatively untried radio component designs were considered and discarded; only tried and proved techniques were used, unless necessity dictated otherwise.

Specificational Requirements

The interchannel crosstalk requirements specified for the radio relay system were based upon meeting the Signal Corps communication system standards for via trunk performance. These standards require that the noise-plus-crosstalk power at the zero level point of any voice frequency channel in via trunk use not exceed 38 decibels (db) adjusted with *F1A* weighting. The interchannel crosstalk performance specification of the radio relay system was based upon meeting this requirement over an 8-jump system. Interchannel crosstalk was specified on the basis of random noise loading as illustrated in Fig. 1. Random noise, restricted to the 17.5 to 56-kc range at a mean power level of 3 db above 1 milliwatt, and a simulated carrier pilot tone at a frequency of 68 kc and at a level of 0 db above 1 milliwatt were used to load the radio relay system at full deviation, with an over-all circuit loss of 0 db on a modu-

Table I. Noise Modulation Crosstalk Tests with Temperature and Humidity Variations

Frequency, Mc	−40 Degrees Fahrenheit		+78 Degrees Fahrenheit		+132 Degrees Fahrenheit	
	C/A	C/B	C/A	C/B	C/A	C/B
	Temperature Test					
100.25	47	51	54	54	54	54
224.75	50	53	48	53	47	52
225.5	51	48.5	47	53	53	51
399.5	50	53	53.5	53.5	52	54
	Standard Signals Corps Humidity Test					
Frequency, Mc	1 Cycle		2 Cycles		5 Cycles	
	C/A	C/B	C/A	C/B	C/A	C/B
224.75	53	46	52	50.5	56	55

Note: All ratios shown are in db.

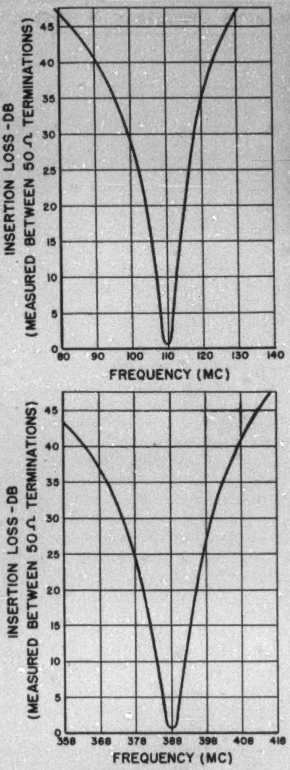


Fig. 2. Insertion loss characteristics of antenna band-pass filters. Insertion loss at tuned frequency less than 1 db

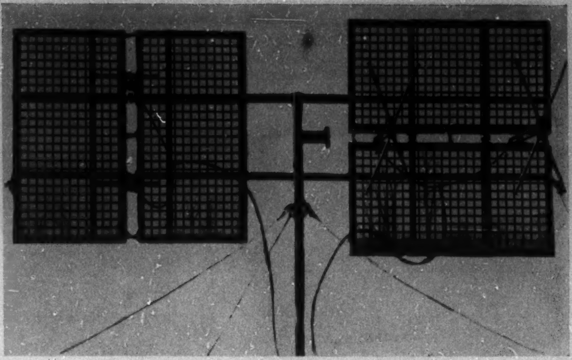


Fig. 3. Antenna assemblies used with radio set AN/TRC-24

lation frequency basis. The mean noise power in the 12 to 16-kc, 32 to 36-kc, and 60 to 64-kc ranges, designated as the *A*, *C*, and *B* channels respectively in Fig. 1, were then measured at the receiver output. The ratio of the in band to out-of-band power levels, namely the C/A and C/B ratios, were specified to be not less than 45 db. Table I lists the data obtained when the radio relay system was subjected to extremes of temperature and humidity variations at various operating radio frequencies. These data indicate not only compliance with the specification but are evidence of the performance that may be expected under severe field conditions. The random noise method of crosstalk measurement has been used throughout the development and production phases.

The range of operating frequencies selected for this equipment was from 50 megacycles (mc) to 600 mc. The lower limit of 50 mc was chosen since it is the approximate lowest frequency not subject to sky wave transmission, while the upper frequency of 600 mc was dictated because of the nonavailability of suitable transmitting and receiving tubes for higher frequency operation. It was decided that the 50 to 100-mc and the 400 to 600-mc ranges would not be included in the basic contract since the 50 to 100-mc spectrum was already in use by other radio relay and tactical equipment, and since the 400 to 600-mc range presented a development problem of considerable magnitude. The frequency range of 100 to 400 mc was accordingly specified. The design in the 100 to 400-mc components was based, however, on ultimately achieving operation in the 50 to 100- and 400 to 600-mc ranges at minimum additional cost. It appeared that this could best be done with basic transmitting and receiving units with the operating radio frequencies obtained by means of physically interchangeable plug-in units. The range of 100 to 400 mc was considered to be best provided in two ranges with each of the plug-in units providing about an octave coverage. Results of studies had indicated that this coverage was the maximum permissible consistent with good electrical performance and equipment simplicity.

The components allied to a particular frequency range of operation have been so designed and packaged that, if operation is restricted to one frequency range, only those components which are applicable to that frequency band need be procured or transported to the operating site. A considerable saving in money and logistic complication is thereby effected.

Operational and Maintenance Considerations

The design of all tactical radio equipment is based upon meeting the particular functional requirements under extreme temperature and humidity and rough field use. With the short-range single channel equipments, emphasis is placed on small size, light weight, submersion proofing, and "flick" tuning. Since radio relay equipment is continuously operated and used more to the rear areas, extreme emphasis is not placed on size, weight, and speed of changing frequencies, but ease of maintenance and reliability are stressed. The tactical short-range equipments are almost always operated by relatively unskilled personnel and these equipments are characterized by a minimum of controls and adjustments. The radio relay philosophy is that, even though the operator may not be trained to maintain or repair the equipment, sufficient controls, metering positions, and alarms are included to signify performance on a system and individual component basis so that a system may be adequately aligned and monitored and the quality of transmission determined. Normally, repair to radio relay components is carried out at depots or in transportable maintenance vehicles where skilled personnel and adequate test equipment are available. To facilitate maintenance, great pains were taken to allow for open, orderly construction with removable, accessible subassemblies that could be removed, tested, and interchanged in their entirety. Every precaution was taken to

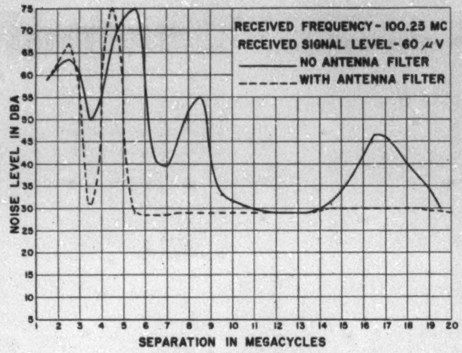


Fig. 4. Data showing reduction in transmitter-receiver frequency spacing by use of antenna filter

NOISE PLUS CROSSTALK IN VOICE CHANNELS-DBA												
ORDER WIRE	1	2	3	4	5	6	7	8	9	10	11	12
				NOT OPERATING								
☒	☒	☒	☒		☒	23 (16)	20 (16)	20 (16)	22 (15)	27 (16)	22 (17)	23 (18)
☒	27 (19)	25 (20)	24 (17)		☒	☒	☒	☒	☒	20 (15)	27 (17)	25 (17)
☒	30 (17)	32 (17)	32 (17)		37 (15)	32 (15)	29 (15)	☒	☒	☒	☒	☒
☒	☒	☒	32 (16)		37 (15)	33 (15)	32 (15)	19 (15)	26 (15)	☒	☒	☒

Fig. 5. Interchannel crosstalk tests over 2-jump 12-channel radio relay system with speech loading

have all components work below their permissible ratings to assure reliable performance under continuous service with extremes of temperature and humidity.

Reliable military use in geographically isolated areas requires that the equipment be issued with all running spares, alignment tools, and other special items. Sufficient running spare parts and operator maintenance equipment are packaged and provided with each radio relay terminal or repeater station to assure 3 months' continuous operation.

In so far as possible, the components of the radio relay system will be maintained with standard items of military test equipment. Two specific developments have been carried out to provide the necessary test instruments that were not available as standard items at the time development of the radio relay system was initiated. The first of these is signal generator *SG-3*, built by Measurements Corporation, which provides frequency-modulated signals over the 50 to 400-mc frequency range. The basic problems associated with this development were the low distortion frequency-modulated performance required and the 8-to-1 frequency coverage. The second item, especially developed by M. C. Jones Company, was radio frequency wattmeter *ME-82*. This wattmeter is usable for measurement and continuous dissipation of radio frequency power at a level of 100 watts in the frequency range of 50 to 600 mc.

Table II. Interchannel Crosstalk Ratios Obtained with Noise Modulation Over a 1-Jump and a 2-Jump Radio Relay System

Number of Jumps	Crosstalk Ratios, Db	
	C/A	C/B
1	48	51
2	44.5	48

Quite often the repeater and terminal stations of a radio relay system are installed simultaneously or in random sequence. It was therefore necessary to include in the design of the equipment means to permit the operator to align completely his station prior to the reception of a remote signal. It is also probable that a complete radio system will be installed prior to the interconnection of the radio terminal station with the telephone terminal equipment and cable. Means have been provided to permit complete alignment of the radio relay system independent of the wire terminal equipment and cable circuits to assure complete readiness of the radio system prior to the connection of the telephone terminal equipment so that completion of alignment of the entire system is not so complicated and time consuming.

To obtain proper system supervision and maintenance, order wire facilities were provided that permitted signaling and talking between any radio terminal station, radio relay station, attended telephone repeater station, or telephone terminal station on a party-line basis. It is also possible to communicate with any of these stations from either half, or between halves, of a repeater station whose easterly and westerly components may be separated by a length of spiral-4 cable, due, e.g., to siting considerations.

Antenna System Planning and Design

The design of the *AN/TRC-24*, incorporating greatly extended frequency range, increased traffic capacity, and operation in the ultrahigh-frequency range, led inevitably to greater weight and size than earlier very-high-frequency sets. To minimize this disadvantage, every effort was made to eliminate nonessentials. The *AN/TRC-24*, as used in the field, is only 20 per cent more in volume and 18 per cent more in weight than previous 4-voice channel very high-frequency equipment, which tuned only narrow frequency ranges.

A major saving in weight and volume was obtained by a new design of the antenna system and its associated mast. Previous military very high-frequency systems have used separate masts for each transmitting or receiving array; the masts have represented a large fraction of the bulk of each station. The feasibility of mounting transmitting and receiving antennas on the same mast without exceeding the size dictated by tactical considerations was determined at SCEL by the design of a wide-band "fat-dipole" antenna and corner reflector, used in conjunction with a high-pass low-pass branching filter which required no tuning adjustments over the range of

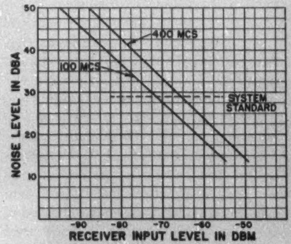


Fig. 6. System quieting characteristic

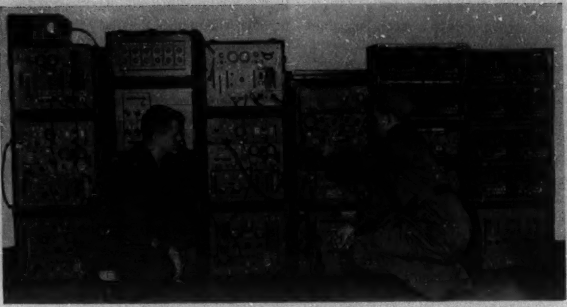


Fig. 7. 12-channel radio relay terminal station with spare components

100 to 225 mc.[2] Though this design possessed the attractions of operational simplicity and relatively low cost, the design was subsequently modified during the development phase to provide 12 tunable band-pass filters covering the 100 to 400-mc range. These filters are used between the antenna and transmitter or antenna and receiver, and provide enough decoupling between the transmitting and receiving antennas to allow a small transmitter to receiver frequency spacings when only one mast is used. The band-pass filters are quite effective in reducing spurious emissions of the transmitter and spurious responses of the receiver, providing needed flexibility in frequency assignments used in field systems. Typical attenuation versus frequency curves are shown in Fig. 2 for the highest and lowest frequency antenna filters.

The antenna design finally selected, shown in Fig. 3, uses a common reflector for the 100 to 400-mc range, and may be operated either horizontally or vertically polarized. In special situations separate masts may be used for transmitting and receiving antennas.

Test and Evaluation at SCEL

Preproduction models of radio set *AN/TRC-24* were delivered to the SCEL in November 1952. Approximately 5 months were spent in evaluation of the equipment by laboratory testing, temperature, humidity, shock, vibration and bounce tests, and by field and systems testing. Electrical specifications were found to be met with only minor exceptions, and the design of the equipment was found to be exceptionally good mechanically. The contractor wisely placed heavy emphasis on mechanical engineering from the inception of his effort. No difficulties were encountered in the interoperation of the radio, cable, and carrier systems. Though only short systems were tested because of lack of sufficient equipment, operation of simulated tactical circuits, with military personnel, gave satisfactory results. An evaluation of the reliability and quality of the system after extended periods of operation was not possible, but is in progress at SCEL now. A brief summary of some of the most interesting test results is given in what follows.

A series of tests was made at the SCEL to determine minimum transmitter to receiver frequency spacings permissible both with and without the antenna filters described in the foregoing. Fig. 4 illustrates the results obtained with a received radio-frequency carrier frequency of 100.25 mc while the interfering source was varied in frequency. The interfering source was a transmitter operating at a normal power output level of 80 watts with noise modulation on the carrier simulating the effect of multichannel speech loading. The disturbing effect was noted by measuring the ambient noise in a voice channel of the 12-channel telephone terminal equipment which terminated the receiver output. The noise levels obtained, which are plotted versus the separation in mc between the received signal and the interfering source, were measured at the 0 db system level point with an *F1A* weighted characteristic. The measured noise level with the interfering transmitter turned off was 28 db adjusted. The measured cross coupling loss between the transmitting, or interfering source, antenna and the receiving antenna was 26 db.

It was realized at the time the development specification was prepared that the wide tuning range specified would result in some sacrifices in systems characteristics. One of the most notable characteristics affected was the sensitivity of the receiver. The noise performance of the receiver as specified was based on only small margin over the minimum performance indicated by theoretical calculations, but due concessions were made to the unavoidable sacrifice in noise figure dictated by the wide tuning ranges required. The measured noise figures at the extremes of temperature and also with high relative humidity do not exceed 12 db in the 100- to 225-mc range and do not exceed 15 db in the 225 to 400-mc range.

Systems tests were made over a 2-jump radio relay circuit terminated with the 12-channel carrier telephone terminal to correlate the interchannel cross-talk performance measured with noise loading techniques with that obtained under speech loading. Six channels were simultaneously loaded with speech modulation at a level of approximately −2 volume units per channel at the system 0 level

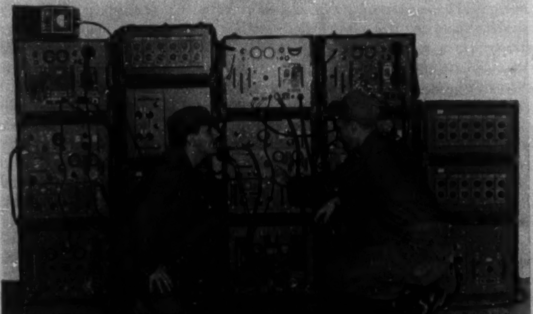


Fig. 8. 12-channel radio relay repeater station with spare components

point. This degree of loading, which had been computed as being proper to simulate expected field traffic, had formed the basis for the determination of the peak system load with the noise modulation method of measuring interchannel crosstalk. Fig. 5 illustrates the results obtained over the 2-jump circuit. The X marks indicate those channels which were modulated simultaneously. The readings of idle noise-plus-crosstalk power which were measured in the unmodulated channels are noted under each measured channel. The readings in parentheses indicate the idle noise power measured with no modulation. All readings were taken at the 0 db system level point with an *F1A* weighted instrument. The military communication system standards require that the noise-plus-crosstalk power in a 4-kc trunk circuit of a 2-jump system be less than 32 db adjusted. The data in Fig. 5 indicate that the radio relay system provides adequate crosstalk performance for via trunk use.

This 2-jump circuit was also tested with the random noise modulation method which has previously been described. The data obtained are shown in Table II with the *C/A* and *C/B* ratios corresponding to the ratios of power measured in the spectra illustrated in Fig. 1, at the system 0 level point. The military communication system standard for 2-jump radio relay system of noise-plus-crosstalk power less than 32 db adjusted in any voice channel corresponds to a mean long-time speech to mean noise-plus-crosstalk power ratio of approximately 47 db or greater. The data obtained therefore indicate that the random noise loading used for test purposes is realistic in simulating expected field performance. Table II indicates the crosstalk products in the 2-jump system add on a power basis, as would random noise. Whether any in-phase addition in long systems will be encountered is not yet known.

Fig. 6 illustrates the noise power measured at the 0 level point of the system of a 12-channel system with varying radio-frequency input power levels to the receiver, at differing radio frequencies. The line drawn parallel to the abscissa indicates the input power level to the receiver required to obtain a noise level of 29 db adjusted, the noise power which is the maximum permissible in any jump of an 8-jump system.

Field Issue and Use

To minimize circuit outages, the radio relay equipment is issued to the using organizations with 100 per cent of the major components in stand-by at the radio terminal stations and with 50 per cent of the major components in stand-by at the radio relay stations. The fact that these stations are frequently located in isolated high points on the terrain, far from maintenance and supply depots, justifies the duplication of components which comprise a large percentage of the total expense, cost, and weight. The radio relay equipment is currently in quantity production by the Western Electric Company and initial shipments have been made to using Army organizations.

Fig. 7 illustrates an operating 12-channel terminal station, complete with stand-by major radio relay components and also a 12-channel carrier telephone terminal with voice frequency ringers for each channel. The stand-by transmitter, transmitter power supply, and receiver are to the extreme left of this figure.

Fig. 8 shows an operating 12-channel repeater station, complete with stand-by major radio relay components. The stand-by transmitter, transmitter power supply, and receiver are to the extreme left of this figure.

References

1. A NEW ULTRAHIGH-FREQUENCY MULTICHANNEL MILITARY RADIO RELAY SYSTEM, **John G. Nordahl.** *AIEE Transactions*, vol. 73, pt. I, Nov. 1954, pp. 526–31.

2. BROAD BAND UNIDIRECTIONAL ANTENNA, **R. Guenther, V. J. Colaguori.** *Technical Memorandum No. M-1317* Signal Corps Engineering Laboratories. Fort Monmouth, N. J., Aug. 1, 1950.

◆

No Discussion

Rural Automatization in the Swedish Telephone System

BERTIL BJUREL
NONMEMBER AIEE

WHEN AN automatization of the Swedish telephone system came under serious consideration at the beginning of the 1920's, there were available two quite distinct types of selectors. One of these selectors was the L. M. Ericsson machine-driven 500-line selector and the other was the Swedish crossbar selector. Telephone plants under consideration at that time for automatization chiefly comprised large telephone plants in densely built-up areas. It was a difficult matter in those days to make a choice between the two types of selectors. One arrangement which would enable the crossbar selector to compete economically with the 500-line selector for large telephone plants was the employment of marker control. Thus marker systems were proposed as early as the beginning of the 1920's by advocates of the crossbar selector. The Board of Telecommunications, however, decided on several grounds not to adopt such systems at that time. One reason was the possible service risks of common control systems which were not sufficiently eliminated until later. Another reason was that in the 1920's it was not possible to determine the magnitude of the decrease in first cost, plausible in itself, that might be obtained with marker systems. Determination of that point required a mathematical treatment that was not developed until later. The choice was therefore given to the L. M. Ericsson machine-driven 500-line selector for the automatization of the large exchanges in question.

Nevertheless, interest in the crossbar selector was great in Sweden at that time. This selector, the idea of which was partially indicated in an application for patent filed in the United States in 1901 by Homer J. Roberts[1] and later more completely developed by J. N. Reynolds in 1913,[2] had by 1919 been developed by the Swedish engineers, G. A. Betulander and N. Palmgren, into a design[3] that was practicable and an improvement in essential respects. The selector had been tried that same year in Stockholm where a trial plant had been constructed on the marker principle. Thus experience with the crossbar selector was already available at

---

Paper 54-424, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 28, 1954; made available for printing August 26, 1954.

BERTIL BJUREL is with the Swedish Board of Telecommunications, Stockholm, Sweden.

the beginning of the 1920's, and in 1926 a crossbar exchange for 3,500 subscribers was completed at Sundsvall. For reasons already stated, this plant was designed as a step-by-step system, which in respect to grouping of devices and routing plan most nearly resembled the well-known Strowger system.

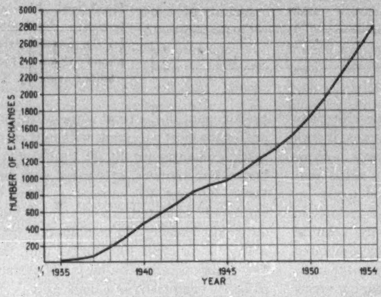


Fig. 1. Increase in number of automatic exchanges 1935–1954

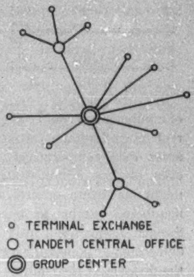


Fig. 2. Principle of layout of the network in a number area

Choice of Selector Type and Size of Exchange for Rural Automatization

When automatization of Swedish rural areas came into question at the beginning of the 1930's comprehensive experience concerning the properties of the crossbar selector was available from the plant at Sundsvall. In view of that selector's great operating reliability and low maintenance cost, the Board of Telecommunications decided to employ the crossbar selector for the automatization of telephone plants in rural areas.

A first system suited to the automatization of rural telephone plants was developed in the years 1931–1932 and an experimental plant for 20 subscribers was put into operation in 1932. On the basis of experience with this plant, an initial series totalling 45 exchanges was designed for unit construction and manufactured for 10, 20, 50, and 90 subscribers. These first unit automatic exchanges were made with single contacts for relays and crossbar selectors. Twin contacts were adopted in 1935 and a new, somewhat modified type was built in a series of 55 exchanges. The difficulty in procuring manual operators for the small telephone plants of rural areas, and the advantages derived by subscribers from automatization, such as 24 hour service, led to ever-increasing interest in automatization of the rural service. Following further improvements, and after far-reaching standardization of the unit automatic exchanges had been carried through, there were produced up to the middle of the 1940's no fewer than 1,200 crossbar plants for the automatization of rural areas. These plants were for the most part of unit construction made in sizes for the connection of 10, 20, 50, and 100 subscribers and could be connected to parent automatic or manual telephone plants.

In view of advances in telephone technique in the 1930's and the extensive experience that had been gained from crossbar selector plants for small telephone exchanges, at the beginning of the 1940's the time had come for further redesigning of the rural unit automatic exchanges to comprise extensive standardization and improvement of operating reliability. Redesigning was carried out; after thorough economic studies the sizes for the unit automatic exchanges were fixed at 40, 60, and 100 numbers, and the new types were given the designation of Standard 41.[4]

In all these types of unit automatic exchanges for the automatization of rural areas, the crossbar selector was employed in principle as a directly or step-by-step operated selector. Interest in marker systems was stimulated during the 1940's by the experience gained in the Bell System crossbar offices in the United States. Work on the development of marker systems for medium and large telephone plants was started about this time and resulted about the mid-1940's in the development of a Swedish marker system known as *A-204*.[5] Exchanges of marker type were introduced as parent exchanges for rural unit automatic exchanges and for all local exchanges larger than 200 subscribers. For exchanges of less than 200 numbers the marker system was not found to display any economic or technical advantages over the existing step-by-step system. Consequently the step-by-step system, Standard *41*, was retained for the actual rural automatization as the system best suited economically and technically to Swedish conditions. To provide for the need of exchanges between 100 and 200 numbers auxiliary equipment was developed enabling a 100-number unit automatic exchange to be combined with a 40-, 60-, or another 100-number unit so as to obtain exchange capacities of 140, 160, and 200 numbers.

Since the start of automatization about 2,800 automatic telephone exchanges have been installed in Sweden; see Fig. 1. Of these some 2,100 have been automatized with the Administration's own standard unit automatic exchanges. At the present time about 70 per cent of the country's 2,000,000 telephones are connected to automatic telephone exchanges. The remaining telephones not yet automatized, amounting to 30 per cent or 600,000 telephones, are distributed over no fewer than 4,600 separate telephone exchanges. Thus it is seen that a typical feature of the Swedish telephone system is that the exchanges in the majority of cases are small. The exchanges are comparatively scattered, resulting in long journeys for the maintenance personnel. For this reason the exchanges must be constructed with the idea of the greatest possible operating reliability and the least possible maintenance.

Numbering Scheme

To enable full automatization of traffic between all telephone subscribers of the country to be progressively carried through, a numbering scheme suitable for the purpose was drawn up at an early stage. By this scheme, the whole country is divided into a series of number areas in each of which all subscribers are allotted a separate directory number. A subscriber's directory number contains 5 or 6 digits; the first two usually represent the code of the particular exchange within the number area. There are about 300 number areas, known as net-work groups in Sweden. Each number area has been allotted a separate area code number, network group number. For all telephone traffic between different number areas the called areas' code number is dialed first and then the 5- or 6-digit subscriber's number is dialed. The area code numbers contain 3 or 4 digits, and the digit 0 is always used as first digit. The sizes of the number areas are governed by traffic considerations and by technical and economic considerations. From the subscriber's point of view it would be advantageous if the boundaries

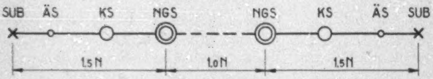


Fig. 3 (left). New attenuation distribution for the Swedish telephone network. Notations as in Fig. 2

of number areas could coincide with easily remembered administrative areas, if such areas were comparatively large. However, obvious administrative areas that might constitute natural and easily remembered number areas are lacking in Sweden, and consequently only technical and economic points of view have dictated the division of the country into the comparatively large number of number areas.

Network Layout

To enable telephone traffic to be handled economically the circuit network, representing the largest part of the investment, must be given an efficient layout. Since the utilization of the circuits increases with increasing circuit group size, the network layout should be aimed at forming large circuit groups until reaching an economic optimum for the total costs of circuits and switching equipment.

In the different number areas it has been found economically advisable to connect rural subscribers to small local exchanges, called terminal exchanges, and assemble the traffic from these exchanges to a central exchange common to the number area and known as a group center. In addition it has been economically justifiable to connect a small type of parent exchange, known as a tandem central office, between the small terminal exchanges and the group center. As a result the network layout in Sweden has acquired the appearance shown in Fig. 2.

The terminal exchanges, which are predominant in the rural automatization, are for the most part made with unit automatic exchanges for 40, 60, or 100 numbers. For terminal exchanges with between 100 and 200 subscribers there is employed, as stated, a 100-number unit automatic exchange in combination with a 40-, 60-, or other 100-number unit, the two units operating by means of certain auxiliary equipment as a single unit.

Table I. Number of Markings on Subscriber's Meter for Different Types of Calls

Calls Between	Number of Markings on Meter
Subscribers at same exchange..	one step on meter independent of length of call
Subscribers at different exchanges in same number area..	one step on meter for each 6-minute period started
Subscribers in different number areas:..	one step on meter for each:
Distances less than 45 kilometers..	45-second period started
Distances of 45 to 90 kilometers..	24-second period started
Distances of 90 to 180 kilometers..	15- second period started
Distances of 180 to 270 kilometers*..	12-second period started
Distances of 270 to 450 kilometers..	10-second period started
Distances greater than 450 km..	8-second period started

* Distances greater than 450 kilometers have not yet been converted to subscriber dialing.

The network layout in the different number areas is star-shaped, as may be seen in Fig. 2. Nevertheless, if the traffic is large enough, in some cases direct routes may be arranged between two tandem exchanges in one number area. In such cases the traffic may be handled in the same way as in the system applied in the United States with alternative routing of traffic. The different routes in the number areas are given dimensions appropriate for 1 to 2-per-cent grade of service in accordance with Erlang's loss formula.

As regards traffic between the 300 number areas of the country, a network is projected including transit centers and arranged for carrying the traffic also by alternative routing. The transit centers will be constructed with crossbar selectors and provide automatic compensation of attenuation. Because of the small geographic extent of the number areas the insertion of repeaters at the different exchanges is not required in these areas. Therefore, all these exchanges, i.e., terminal exchanges, tandem exchanges, and network group exchanges, are made with selectors for 2-wire circuits.

Distribution of Attenuation

Properly balanced distribution of attenuation is of the greatest importance for the economy of a telephone system, and this in turn is closely connected with the type of telephone instrument used. Formerly the telephone instrument used in Sweden had a reference equivalent for transmission of 0.8±0.4 neper and for reception of 0.1±0.2 neper. While this older instrument was in use the Comité Consultatif International Telephonique maximum value of 4.6 nepers was applied in Sweden as an over-all reference equivalent in the telephone system. During 1954 a new telephone instrument was developed showing a reference equivalent for transmission of 0.3±0.2 neper and for reception of −0.5±0.2 neper. It conjunction with the adoption of this improved telephone instrument, it was possible to raise the tolerated line attenuation in each number area by about 0.4 neper over what was previously permitted, and the total reference equivalent for telephone calls was progressively reduced to 4.0 nepers. At present 1.5-neper line and exchange attenuation is permitted from the number area's central exchange, the network group exchange, to a subscriber in the area, including microphone feeding current attenuation on transmission. By the insertion of the necessary repeaters there is obtained between the central exchanges of the different number areas an attenuation which may not exceed 1.0 neper. Thus the distribution of attenuation with the introduction of the new telephone instrument will have the appearance shown in Fig. 3. There are no fixed sectional attenuation values whatever in the number areas for the circuits between the different exchanges and the subscriber lines. How the tolerated attenuation of 1.5 nepers may be utilized in the most economical manner is determined by economic investigation in each separate case. The new attenuation standards have led to extremely large savings in the construction of junction circuits and subscriber circuit networks.

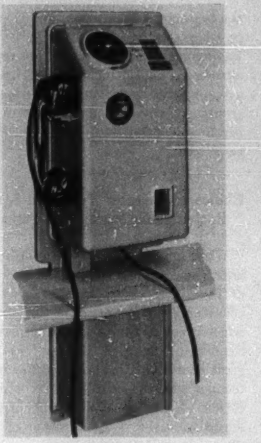


Fig. 4. Automatic call box instrument for three types of coins

Telephone Rates for Rural Automatization

By means of extensive standardization of telephone equipments and standardization of the work in the entire telephone field, it has been possible in conjunction with the automatization of the country's telephone traffic to reduce the costs for handling the traffic. The Swedish Board of Telecommunications has mainly let the subscribers benefit by the reduction in the form of decreased telephone rates and improved service. Because the structure of the rates is closely related to the layout of the technical equipment and exercises an appreciable influence on the volume of traffic, a brief account of the rates applied and the methods of charging with automatic telephone operation will be given here.

On signing a contract for telephone subscription a subscriber in the Swedish system pays an installation charge of 200 kronor. The amount of installation charge is in no way dependent on the costs to the Administration for connection of the individual subscriber, and in sparsely inhabited regions these costs may amount to considerable sums. In the automatized networks the subscriber usually pays in addition a fixed charge amounting to 86 kronor per annum. This charge applies to subscribers with their own direct lines to the telephone exchange. For subscribers who share a telephone line the annual cost is 20 kronor less, i.e., 66 kronor. Party lines occur to an extremely limited extent in automatized networks and the maximum number of subscribers on one and the same line in an automatic network amounts to only two. At present the number of subscribers with shared lines throughout the country is only 2.5 per cent of the number of subscribers and of these about three fourths are connected to manual telephone plants.

The choice in Sweden for charging of the costs for the individual telephone calls lay between systems for automatic printing of call tickets or systems for debiting the calls on the normal subscribers' meters. The latter system had been employed successfully abroad and it undoubtedly represented the most favorable method from the technical and economic viewpoints for the Swedish telephone service. The Board of Telecommunications therefore decided on this charging method at an early stage and it has been successfully in use several years both for local calls and for automatic toll calls. Each recording on a subscriber's meter is at present charged to the subscriber at the rate of 6 öre; the number of steps recorded in the meter for different types of calls are listed in Table I.

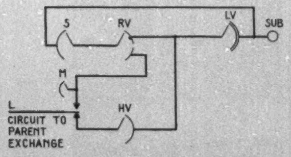


Fig. 5. Trunking arrangement for unit automatic exchange

Each number area normally comprises a population group which has its business interests in the neighborhood of the group center, and this is one of the reasons it was considered warranted to set the rate for all traffic within the number area as low as one step on the meter for each started call period of 6 minutes. Sometimes two neighboring number groups have such strong common interests that benefit of the rate in question is extended to comprise all traffic between subscribers in two neighboring number areas. The jump from 6-minute periods to 45-second periods is comparatively large. Thus, for subscribers located in the boundary districts between two number areas a special buffer rate is often applied which makes calls cost less than the rate with 45-second periods.

The technical devices for determining the meter time intervals are always located at the tandem exchanges or the network group center exchanges. The correct time interval for the metering is determined by the register employed for establishing the connection. For this purpose the register operates a special charging unit which is normally connected to the circuit equipment for the circuit from the tandem exchange. This unit transmits metering impulses to the calling subscriber's meter. If the call is originated at a terminal exchange metering signals are transmitted by the charging unit at the parent exchange over the occupied circuit to the terminal exchange. A central location for the rate-fixing equipment was considered valuable from the maintenance, standpoint, and no difficulty has arisen in sending metering signals over the occupied circuit while a call is proceeding.

Subscriber's reaction to the introduction of charging for toll traffic on the ordinary subscribers' meters has been favorable. However, certain technical equipment for determination of the cost for separate calls has been requested by subscribers. The need for such equipment has been advanced chiefly by hotels, restaurants, shops, and factories, though even purely private subscribers have in exceptional cases asked for such devices. Such equipment normally consists of meters installed in direct connection with the telephone instruments or private branch exchanges. In some cases automatic private branch exchanges have been provided with control devices, blocking automatic toll traffic from the various extensions so that these are compelled to order such calls with the operator of the private branch exchange.

To facilitate telephoning from public call offices in the automatic telephone network an automatic public telephone instrument for three denominations of coins has been designed; see Fig. 4. Toll calls can be obtained automatically with this instrument and charges for the calls are collected without any manual intervention. A call may be started from the instrument immediately after an initial charge of 20 öre has been inserted and the call may continue for an unlimited period, provided the necessary coins are inserted while the call is going on. This instrument will be produced on a large scale in the next year or two.

Technical Build-up of the Rural Unit Automatic Exchanges

Trunking Arrangement

The trunking arrangement for rural unit automatic exchanges of sizes 40, 60, and 100 numbers is shown in Fig. 5. The subscribers to these are connected to finder switch *S* and to final selector *LV*. Between these selector stages an extra selector stage, discriminating selector *RV* is inserted over which outgoing calls may be connected either to an idle rural line for outgoing traffic or to the final selector in the caller's own exchange. A preselecting stage with preselectors *HV* carries

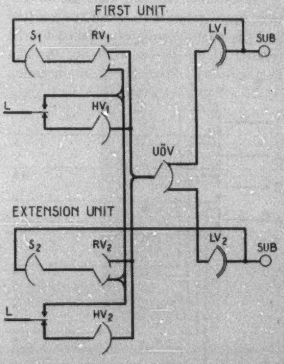


Fig. 6. Trunking arrangement for a rural exchange consisting of two combined unit exchanges

Table II. Number of Switches and Circuits for Unit Automatic Exchanges

Type	Number of Switching Devices				Maximum Number of Circuits		
					All One Way		All Two Way
	S	RV	LV	HV	Incoming	Outgoing	
40-number units	5	5	5	4	4	4	4
60-number units	7	7	7	5	5	6	5
100-number units	9	9	9	7	7	7	7

out connection of incoming traffic to the final selector of the exchange.

Since all terminal exchanges are designed without registers, the outgoing calls must all be directed in the first place to a parent exchange with registers. This connection is done over the nonnumerical finder stage *S* and the lower multiple field of discriminating selector *RV;* see Fig. 5. On connection to a register at the parent exchange the subscriber receives a dial tone and then dials the wanted subscriber's number. The number is registered in the register at the parent exchange and at the same time the selector-repeater *M* at the terminal exchange notes whether the two first digits correspond to its own or another exchange's number.

For an internal call the selector-repeater *M*, normally after receiving three digits, transmits a switchover signal to the discriminating selector *RV* and this latter, following a nonnumerical hunt, connects the subscriber to a free final selector *LV*, the line to the parent exchange being released at the same time. Hunting for a free final selector *LV* is done in the interval between the third and fourth dialed digits. The last two digits in the subscriber number directly set the final selector to the called subscriber.

For external outgoing calls the switchover signal referred to is not transmitted and the connection is established by the register at the parent exchange through to the called subscriber. On incoming calls the preselector *HV* is started and establishes connection to a free final selector *LV*. The operation of the final selector then proceeds in the same way as for internal traffic.

If all lines to the parent exchange should be engaged by outgoing calls, the subscriber is connected over the submultiple of the discriminating selector *RV* to a special relay group called the blind circuit. This relay group, which is provided with a selector repeater, gives the subscriber dial tone. If the number then dialed refers to an internal call, the blind circuit sends a switching signal to the discriminating selector *RV* and a final selector *LV* is connected in the ordinary way. If, however, the number refers to an outgoing external call, the subscriber receives a busy tone.

Requirements regarding the number of switching devices vary, of course, from exchange to exchange, but to insure effective standardization of the construction of the exchanges all the smaller automatic exchanges, i.e., those for 40, 60, and 100 numbers, are equipped with a fixed number of switching devices. The number of switching devices and circuits for each type of exchange was fixed after extensive traffic investigations, and may be seen in Table II. If in exceptional cases the internal or external traffic is found to be unusually large, there is no choice but to operate the exchange with a number of subscribers smaller than it is normally designed for.

As stated, exchanges with a subscriber stock between 100 and 200 are provided with a 100-number unit automatic exchange in combination with a 40-, 60-, or another 100-number unit. The arrangement for assembling two such units to one large unit is extremely simple. The trunking arrangement for such a combined unit is shown in Fig. 6. Calls incoming to the combined unit from a parent exchange are directed either over the preselector *HV1* in the first unit or over the preselector *HV2* in the extension unit. It is immaterial which of the preselectors is used, this depending solely on the line that is called from the parent exchange. A special extension selector *UÖV* is called over the preselector. Selector *UÖV* consists of a group selector which has two exits with a hunting capacity of ten for each exit. The selector is directed from the parent exchange by the digit 0 to the first unit's final selector *LV1* or by the digit 1 to the extension unit's final selector *LV2*. For purely internal calls the extension selector *UÖV* is operated in the same way. For outgoing traffic the units operate in the way described for a separate unit.

DESIGN AND OPERATION OF SELECTOR DEVICES

A detailed account of the design and operation of the selector devices in the three types of unit exchanges would be outside the scope of this paper. Therefore only some of the most characteristic features of the different devices will be given here.

The bank contact assemblies of the Swedish crossbar selector can normally be made with eight make contacts per contact assembly. The number of conductors for each connected circuit in the rural unit automatic exchanges amounts to only four. It was therefore decided for reasons of economy to make double use of the crossbar selector hold magnet units, a

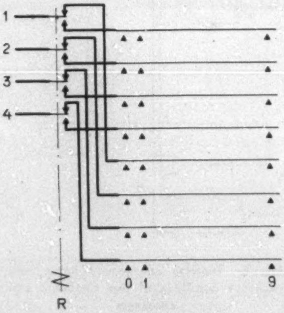


Fig. 7 (left). Principle of double utilization of crossbar selector bridge

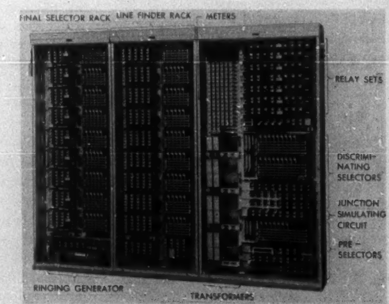


Fig. 8 (right). View of 100-number unit automatic exchange

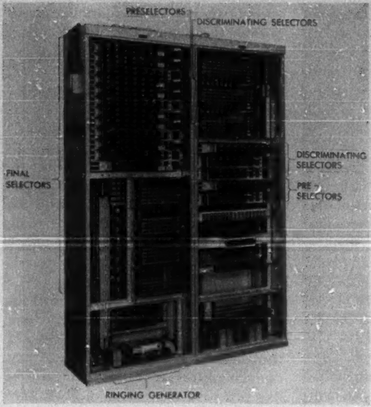


Fig. 9 (left). Front view of 60-number unit automatic exchange

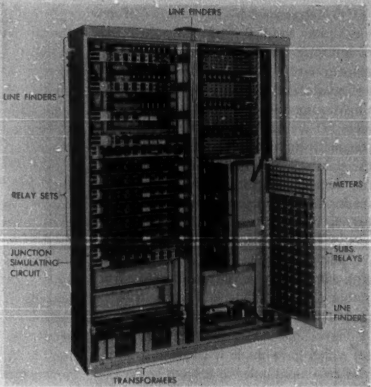


Fig. 10 (right). Back view of 60-number unit automatic exchange

switching relay being used for establishing connection over one or the other of the contact assembly halves of the hold magnet unit. The principle of the double utilization of the hold magnet unit's contact assembly is shown in Fig. 7, where *R* is the switching relay in question. With ten selecting magnets, which is normal for crossbar selectors, each doubly utilized hold magnet unit represents a 4-terminal 20-number selector.

With the stated double utilization of crossbar selector hold magnet units, to make a 100-number finder switch a crossbar selector with five hold magnets is required and for the 100-number standardized exchange unit as a whole, with its nine finder switches, a total of nine 5-hold magnet crossbar selectors is required; see Fig. 8. For a 60-number finder switch three hold magnets per finder switch are required or for the whole 60-number standardized exchange unit with its seven finder switches, $3\times7=21$ hold magnets. As three hold magnet crossbar selectors would be rather small units, the seven finders are made up of three 6-hold magnet crossbar selectors with two finder switches per crossbar selector and one crossbar selector of three hold magnets; see Figs. 9 and 10. Finder switches in the smallest type of standardized exchange unit, which have only a capacity of 40 numbers, require only two hold magnets per finder switch. Thus the five finder switches of the unit take $2\times5=10$ hold magnets; see Figs. 11 and 12. For these five finder switches in this case one 6-hold magnet crossbar selector and one crossbar selector with four hold magnets have been utilized.

On calls from a subscriber the line relay energized directly indicates the contact assemblies to which the calling subscriber is connected. The tens in the subscriber's number, for instance, in a 100-number unit will indicate which bar is to be actuated in a free finder switch. The units digit will indicate which of the five hold magnets in the finder switch is to be used, also whether the subscriber is connected to the first or to the second of the two contact assembly halves of the hold magnet unit in question. With the smaller unit automatic exchanges it is, of course, the units digits that indicate the number of the selecting bar in the finder switch and the tens digits which indicate the hold magnet number and the appropriate one of the contact assembly halves.

Each finder switch has a control relay set of its own and in a unit exchange there is a common control device for all finder switches. To eliminate the risk of a defective finder switch blocking a unit exchange when traffic is low, the finder switches have been provided with call distributors which progressively route the calls to the different finder switches.

The final selectors in the unit exchanges are made up of crossbar selectors arranged in the same way as for finder switches; see Figs. 8–12. Each final selector has a separate control relay set. In the smallest standardized exchange units, where each crossbar selector represents several final selectors, a certain interdependence between the control relay sets of the separate final selectors is inevitable. In the 100-number unit the crossbar selector selecting magnets are set directly by the tens digit and the units digit operate the hold magnets as also the switching relay.

In the two smallest unit exchanges it is the units digit that operates the selecting magnets and the tens digit the hold magnets. As the tens digit is received before the units digit, the value of the tens digit must be stored by special relays during the time it takes to record directly on the final selector selecting magnets, the units digit coming after.

With internal calls the microphone feeding current for both calling and called subscriber is obtained from the control relay sets of the final selectors. With calls incoming from a parent exchange the final selector furnishes microphone feeding current only to the called subscriber. The individual control relay sets of the final selectors contain equipment for the transmission of ringing signals, busy tone, ringing tone, etc.

The hunting capacity of the discriminating selectors *RV* and the preselectors

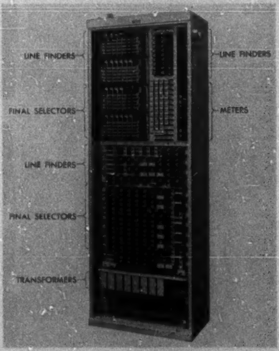


Fig. 11. Front view of 40-number unit automatic exchange

HV is determined by the number of devices to be reached in the different types of units by the selectors concerned. All discriminating selectors *RV* in a unit automatic exchange are controlled by a single common control relay set. In this relay set a relay chain is included which progressively connects testing relays to the final selectors or the testing wires of the outgoing circuits. A common relay set for similar purposes is allotted to the preselectors of the exchange. The relay sets of both the discriminating selectors *RV* and the preselectors *HV* are provided with call distributors to eliminate the risk of the unit being blocked by faulty devices when traffic is low.

The extension selectors *UÖV* in Fig. 13, which have made it possible to combine two small units to make a larger cooperating unit, are made up of three crossbar selectors each with five hold magnets. Each hold magnet unit represents an extension selector and due to the double utilization there is obtained a hunting capacity of ten for each of the two exits.

SIGNALING SYSTEMS FOR RURAL CIRCUITS

Extremely important for reliable operation of the rural network is a reliable system for signal transmission between the different exchanges. Usually a nonphantomizable fully reliable d-c system is therefore employed for the terminal exchanges. However, systems with inductive selection do occur in cases where phantomizing must be done.

In the d-c system signals are sent from the calling exchange in the direction of the called in the form of making and breaking the direct current in the line loop. Signals in the opposite direction are sent by means of reversal of current flow. Digit transmission is thus done by decimal interruption impulses and transmission of call metering signals by means of momentary reversal of current flow.

The signalling systems are so designed that defective lines are automatically indicated by the relay equipments of the lines at their outgoing ends. When a line fault or fault in the switching equipment for the incoming end of the line is indicated, the line is automatically disconnected. The line relay assembly at the outgoing exchange after that periodically sends out fictitious calls, and if the line fault is removed or possibly was of a temporary nature the line is again automatically connected for operation. If the fault persists for any length of time, it is announced to the nearest attended exchange. This equipment for fault indicating of lines, which has been adopted for the whole of the Swedish automatized network, has been found to be an excellent means of reducing maintenance work, particularly at the nonattended telephone plants.

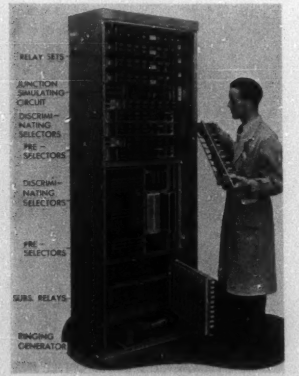


Fig. 12. Back view of 40-number unit automatic exchange

SPECIAL FACILITIES OF INTEREST

Fault Indicating Means

Fault locating in the unit automatic exchanges is facilitated by various marking lamps which roughly give the type of fault and the position in the equipment. In addition the individual devices, are provided for the purpose of testing for faults with a blocking switch, a busy lamp, and a test jack.

Fault Alarm Signalling

In the event of any fault, including faults in exchanges, notification signals are transmitted automatically to the nearest exchange with maintenance personnel. On transmission of the report various codes state whether the fault is of a serious nature and thus requires immediate attention or whether the disturbing effect of the fault is small and attention therefore may be deferred until a suitable occasion to visit the exchange presents itself. The alarm is transmitted over the normal circuits used for traffic at a moment when they are free.

Automatic Answering Circuit

To allow test calls to remote unit automatic exchanges these are provided with an automatic answering circuit connected to the unit in the same way as an ordinary subscriber's line. The answering circuit on call sends an answer signal and a special code tone to the calling lineman or possibly to a centrally located robot.

Vacancy Connection

Vacant subscriber numbers are marked by insertion of a break stick between certain contacts in one of the subscriber's line relays. On calling a vacant subscriber number the calling subscriber receives a special information tone. By calling the nearest information bureau the caller can obtain information regarding the vacant-connected number.

Absent Subscriber's Service

If a subscriber lifts his handset without subsequently dialing any number within a predetermined time of about 30 seconds, the exchange equipment automatically causes a switchover of the subscriber's relays to a particular condition which may be referred to as permanent loop condition and in which all switching equipments engaged by the call are disconnected. This particular condition is restored as soon as the subscriber replaces his handset. On a call to a subscriber having his line relays in the permanent loop condition the calling subscriber receives the information tone. This possibility of special marking of a subscriber number may be utilized by subscribers in cases where they wish certain information conveyed to calling subscribers during a given period. The subscriber first rings up the nearest information bureau and there requests that the operator give the desired message to sub-

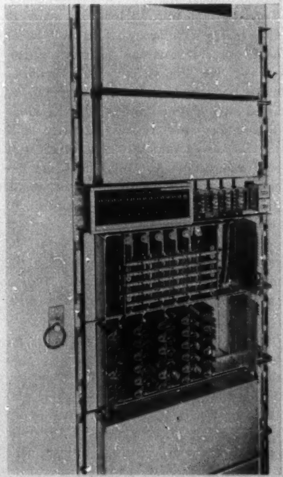


Fig. 13. Rack with extension selector

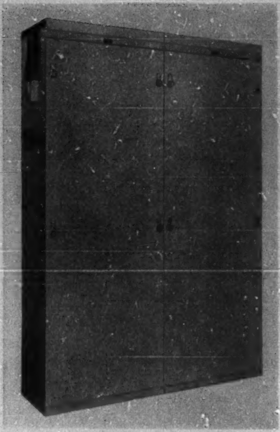


Fig. 14. 60-number unit automatic exchange with wooden casing

scribers who inquire concerning the reason for the information tone. The subscriber then leaves his handset off and the number is automatically connected for reference as soon as the time supervision breaks the call. This facility is also of value for the discovery of short-circuited subscriber lines. In the event an inquiry is received concerning a subscriber number regarding which no communication has been given to the information bureau, generally the line may be considered to be defective or the subscriber may have forgotten to replace his handset.

Congestion Indicating Facilities

For judging the accessibility in the unit automatic exchanges there are congestion meters for finder switches, final selectors, and outgoing toll lines. The unit automatic exchanges are not equipped with any devices for direct measurement of traffic.

Equipment For Time Supervision

Time supervising devices connected to final selectors and junction circuits prevent blocking of the switching equipment in the event of line fault or in conjunction with failure of a clearing signal from one of the subscribers in an established connection.

Power Supply and Heating of the Buildings

The batteries used in the unit automatic exchanges are 36-volt lead accumulators. Voltage variations are comparatively large and for a unit entirely without traffic the voltage may amount to as much as 45 volts. The tolerated voltage variations range from 30 to 45 volts. The batteries are charged by means of mains connected dry rectifiers which are permanently in service. The introduction of rectifier equipment for constant voltage is planned. The exchange rooms are usually heated by electric radiators. These are connected automatically by a temperature regulator if the temperature goes below 4 degrees centigrade.

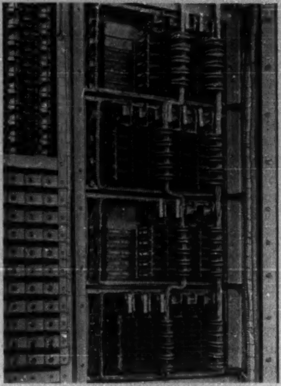


Fig. 15. Bare wiring of finder switch multiples and final selector multiples for 40-number unit

MECHANICAL BUILD-UP OF UNIT AUTOMATIC EXCHANGES

All unit automatic exchanges are constructed as rack units protected against dust by a wooden casing. The fronts and backs of the racks are covered by easily removable masonite covers, as shown in Fig. 14. There are no separate covers for relays and selectors. Where double-sided mounting has been employed for the 40- and 60-number units, the apparatus mounted opposite the crossbar selectors is detachable or can be swung out; see Figs. 10 and 12. The relay sets are normally made as detachable units with plug and jack connection to the rack. This facilitates manufacture, transport, mounting, and maintenance. The racks, Fig. 14, are provided with jacks for the connection to the main distribution frame. By this means the installation of a unit exchange as well as its replacement with another one of different capacity is greatly facilitated.

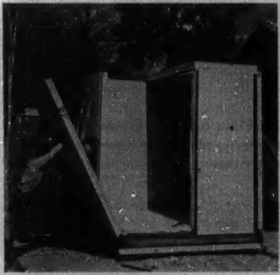


Fig. 16. Erecting standard hut type A1

Wires and cables insulated by polyvinyl chloride are employed for the internal wiring of the unit racks as also for the remaining cabling in the exchanges. In the finder switch multiples and final selector multiples it has been possible to use bare wiring for the 40- and 60-number units; see Fig. 15. This has been possible owing to each crossbar selector constituting more than one finder switch or more than one final selector. This circumstance, which is of some importance from the manufacturing viewpoint, is mentioned because the idea is often held that bare wiring is not possible for finder switches and final selector stages, except in link systems. Bare wiring of course, is used for discriminating selectors and preselectors.

Standardized Types of Exchange Buildings

The rural exchanges are housed in standardized huts which are prefabricated and delivered ready for assembly at the place for the exchange. There are in all three types of huts for all types of ex-

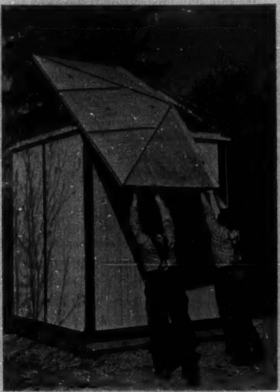


Fig. 17. Erecting standard hut type A1

Fig. 18. Standard hut type A1

Fig. 19. Standard hut type A3

changes, i.e., for 40, 60, 100, 100+40, 100+60, and 100+100 number units. These are type *A1* for maximum 60 number, with internal dimensions: height 2.4 meters, length 2.3 meters, and width 2.0 meters; *A2* for maximum 100 numbers, with internal dimensions: height 2.4 meters, length 3.4 meters, and width 2.0 meters; *A3* for maximum 200 numbers, with dimensions: height 2.4 meters, length 4.0 meters, and width 2.8 meters.

By this standardization it has been possible to keep the cost of the huts down and the time for erection of, say, an *A1* hut amounts to about 5 working days. Figs. 16 and 17 show the erection of an *A1* hut and Figs. 18 and 19 show erected huts *A1* and *A3* respectively. The huts are well heat-insulated by rockwool mats, so that the energy for heating required has been reduced to a minimum.

Standardization of Exchange Assembly

It was found after experience with the assembly of a large number of unit exchanges that it should be possible to decrease the cost of assembly by properly carried-out standardization. Thorough study of various assembly methods has now been completed and new standard methods have been developed. Thus a power distribution panel common to all types of unit exchanges, as seen in Fig. 20, has been designed, this being put up in conjunction with the erection of the hut. For the assembly of the other equipment required for the unit exchanges, other types of panel have been designed, which are manufactured as complete easily mountable units. Of the panels shown in Fig. 20 one contains a rectifier for the power plant, telephone instrument, power distribution blocks, etc. Another panel contains protectors, test jack strips, and cable termination for 200-wire lead-in cable. The number of the last-named panel type is directly determined by the capacity of the exchange. For *A1*-huts one such panel is required as maximum; for *A2*-huts two panels as maximum; for *A3*-huts three panels as maximum. On delivery these cable termination panels are furnished for the line side with one spliced lead-in cable and for the exchange side with cable and plugs for direct connection to the unit racks.

Because of the assembly methods now developed, with extensive preassembly of requisite equipment in the exchange buildings it has been possible to decrease the working time for construction of a unit automatic exchange in the past year as follows:

1. For the 40-number unit from 15 working days to 1 or 2.
2. For the 60-number unit from 16 working days to 2 or 3.
3. For the 100-number unit from 20 working days to 3 or 4.

An attempt has been made to deliver exchange huts assembled and completely mounted with equipment directly from an assembling workshop to the exchange site. However, it is still too early to decide whether this method will be introduced as a general practice.

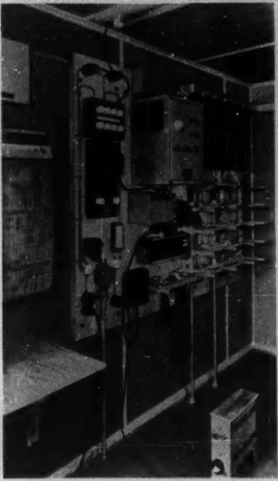


Fig. 20. Different types of standard panels (hut type A2)

Table III. Maintenance Times for Rural Unit Automatic Exchanges

		Number of Hours for Maintenance per Year Including Travelling				
Type of Unit Exchange	Number of Visits per Year for Fault Clearance	For Direct Fault Clearing	For Preventive Maintenance	In Conjunction with Other Visits	Total	Per Connected Local Line (10-Per-Cent Vacancies)
40	5	5.1	23	1	29.1	0.8
60	8	8.3	29	1.5	38.8	0.7
100	13	13	40	2.5	55.5	0.6

Here it may be worth mentioning that very considerable standardization also has been carried out in respect to the manufacture of the telephone equipment. One example of this is the crossbar selector. Whereas 12 years ago it took no fewer than 55 working hours to produce a 10-hold magnet crossbar selector, the same selector can now be produced in only 11 working hours.

Maintenance Costs for Rural Automatic Switchboards

It is extremely difficult to give definite particulars of the maintenance costs for automatic telephone plants, as these costs are governed by a very large number of factors. Thus, in addition to the operating reliability of the exchanges, the ability of the personnel and other local conditions play a great part. Nevertheless efforts have been made to arrive at fairly representative mean values for maintenance work with the Swedish rural exchanges. These values are given in Table III.

Conclusions

1. The crossbar selector has been of the utmost importance in the automatization of telephone offices in the Swedish rural areas. The crossbar selector has made it possible for great operating reliability to

be attained, which is an absolutely necessary requirement with the automatization of unattended telephone exchange.

2. By far-reaching standardization of the type of exchange, huts, and assembly methods the costs of plant can be kept down.

3. As the standardized unit automatic exchanges can be connected to both automatic and manual parent telephone exchanges, it has been possible to select for automatization the plants which, among other things, are the most wanted in view of the difficulty in securing manual operators.

4. Along with the telephone automatization in the rural areas during the last two decades there has occurred a heavy increase in the number of subscribers and in traffic for these areas. This development is due to a large extent to a considerable increase in the standard of living in rural Sweden. Contributing factors also should be the advantages obtained through the automatization and among these in no small degree is the considerable reduction in toll traffic rates accomplished in connection with the automatization.

References

1. TELEPHONE SYSTEM, **Homer J. Roberts.** *United States Patent* 1,297,251, 1901.

2. SELECTOR-SWITCH, **J. N. Reynolds.** *United States Patent 1,139,722*, 1913.

3. SWITCHING DEVICE FOR SELECTING AND OPERATING OF CONTACT GROUPS ARRANGED IN ROWS CROSSING EACH OTHER. **G. A. Betulander, N. Palmgren.** *Swedish Patent 49, 816*, 1919.

4. THE SWEDISH TELEGRAPH ADMINISTRATION'S CROSSBAR SWITCH SYSTEM, STANDARD 41, **Sten Vigren, Artur Finström.** *Tekniska Meddelanden*, Stockholm, Sweden, nos. 1–3 and 6–7, 1944.

5. A NEW COMMON CONTROL AUTOMATIC CROSSBAR TELEPHONE SYSTEM INSTALLED IN SWEDEN, **H. F. Rost, G. Modée, H. Strandlund, F. Ek.** *Electrical Engineering*, vol. 71, Feb. 1952, no. 2, pp. 134–39.

6. TELECOMMUNICATIONS IN SWEDEN—PRESENT AND FUTURE, **Håkan Sterky.** *Tele* (English edition), Stockholm, Sweden, no. 1, 1950.

7. THE CROSSBAR SWITCH—A BRIEF HISTORY. *Ibid.*, no. 2, 1950.

8. TECHNICAL VIEWPOINTS RESPECTING AUTOMATISATION OF TRUNK TRAFFIC, **B. Bjurel, H. O. Björk, E. Waldelius.** *Ibid.*, no. 1, 1953.

No Discussion

Automatic Percussion Welding

A. L. QUINLAN
MEMBER AIEE

PERCUSSION welding is not new. Early work goes back beyond the beginning of the century, but little application has been made of it and only a meager amount of literature is available. However, this method has a real field of usefulness as the application described in this paper will show. The original Vang process wherein a capacitor charged to a high potential, often to several thousand volts, is discharged across the gap between parts as they approach each other under a propelling force, is a good general description of the method used. The arc so produced heats the abutting surfaces before they collide so that a very thin layer of metal is brought to welding temperature. The propelling force, while continuing to act, brings the parts together percussively and the weld is made. Little metal is heated and little heat penetrates the adjoining metal; therefore, the heat balance problem is greatly minimized and different metals weld together with little trouble. There is, however, the problem of protecting personnel from high voltage. Also, the two surfaces being welded must be insulated electrically from each other. This excludes the use of this process for joining the ends of the same piece of metal, as in making a ring.

The project under discussion is the development of a machine for the automatic multiple percussion welding of contact blocks to the ends of an array of small wires extending less than 1/4 inch from molded phenolic plastic. This array of wires, fixed in plastic, forms a part which will herein be called a "comb"; see Fig. 1. When completed, this comb forms the stationary contact member of the Bell System's new wire spring relay. The small blocks of metal on the ends of the wires are cut from a composite tape of which a small portion near the top and/or the bottom surface is palladium, a precious metal. There is a family of combs to be welded, depending on the number and type of contacts required by the code of relays into which each comb is assembled. All top palladium contact surfaces must be located in the same plane across the 12 wire positions of the comb within a tolerance of ±0.002 inch to meet design requirements. In addition, other dimensions for locating the precious metal must be held to close limits for reasons of precious metal economy.

The contact blocks for the wire comb are welded to the wire ends by the automatic percussion welder; see Fig. 2. This is one of the units in an automatic welding and forming line; see Fig. 3. The combs are removed automatically from a magazine and transferred by a reciprocating conveyor into each machine unit in succession. Thus, the clipping, forming, and tinning units necessary for making a complete part are integrated into a line with the percussion welder. This line, by being fully automatic in its operation, is an interesting example of automation.

Reasons for Selection of Percussion Welding

Percussion welding was selected for this process for the following reasons:

1. The electrodes may be placed well away from the weld zone. This is necessary to prevent the electrode clamping force from deflecting the wire ends and causing a misalignment of contacts. The lesser current required for arc welding, as compared to resistance welding, makes it possible to conduct this current through the wires from the opposite end without heating them appreciably.

2. A suitable heat balance in the weld zone can be obtained readily. This would be difficult if the slower butt welding method were used, because of the unequal size and differing electric and heat conductivities of the abutting parts. Slower heating methods loose more heat from conduction through the adjacent metal.

3. The fast welding time recommends its use in high-speed automatic welding machines.

Early Experimental Work

Early experimental work on percussion welding showed that small gas pockets in the weld zone were causing weak welds. Nickel silver was used at that time. Its component of easily volatilized zinc was suspected of causing the trouble. Various metals and alloys were tested and silicon-copper was chosen for the wires in the comb. This alloy, however, was not

---

Paper 54-347, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted May 6, 1954; made available for printing July 20, 1954.

A. L. QUINLAN is with the Western Electric Company, Chicago, Ill.

The author wishes to acknowledge the valued assistance of many co-workers at the Bell Laboratories and at the Western Electric Company in the successful completion of this project and in preparing this paper.

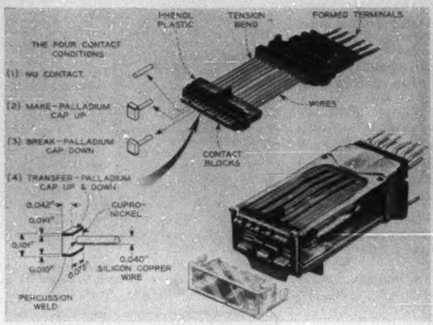

PHENOL PLASTIC
TENSION BEND
FORMED TERMINALS
THE FOUR CONTACT CONDITIONS
(1) NO CONTACT
(2) MAKE—PALLADIUM CAP UP
(3) BREAK—PALLADIUM CAP DOWN
(4) TRANSFER—PALLADIUM CAP UP & DOWN
WIRES
CONTACT BLOCKS
0.042"
0.010"
0.101"
0.010"
0.073"
CUPRO-NICKEL
0.040" SILICON COPPER WIRE
PERCUSSION WELD

Fig. 1. Wire comb with percussion welded contacts and wire spring relay

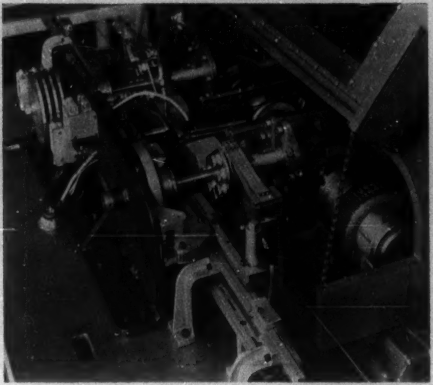


Fig. 2 (right). Percussion welder

suitable for use as the base metal of the contact block, because it was difficult to weld in the roll welding process used for fabrication of the contact tape. A 70–30 per cent cupro-nickel was selected for this base metal because it approximated the resistance of the palladium, a necessary condition for roll welding, and did not have an easily volatile component to weaken the percussion welds.

Another subject investigated was the speed of approach of the parts during the percussion welding operation. The comb and its array of wires are held stationary and the contact block is moved to the wires to make the weld. Various speeds of the contact block carriage or "gun," from approximately 1 to 80 inches per second, were tried. The best welds were obtained by a spring-actuated gun which attained a velocity at impact of approximately 40 inches per second.

The Automatic Welder

The automatic percussion welder contains duplicate welder heads which are mirror images of each other. Each gun may weld a contact each cycle on its own half of the comb; see Fig. 4. After each welding operation or 1 cycle, the comb is indexed to the next pair of wire centers. Six cycles complete the welding of 12 contacts, or a lesser number if required. The guns do not weld at exactly the same time. There is an interval of 1 degree of revolution of the main shaft between the firing points to prevent electrical or mechanical interference. The welder was designed to select the type of tape, cut the contact from it and weld it in 1 cycle, to avoid the handling problems associated with precutting and magazining contact blocks.

TAPE SUPPLY AND SELECTOR

One of four contact conditions will apply for each wire: no contact; contact with palladium cap up; palladium cap down; or palladium cap both up and down. This requires three reels of tape on the right for one head and three on the left for the other. Adjustable knobs on both right and left tape-feed cams are set for one of the four tape-feed conditions for each wire position of the comb. Thereby, any combination of contact conditions can be set up to make parts for any code of relay.

CONTACT SHEAR AND TRANSFER

The three tapes for either head feed from opposite sides. They enter the shearing dies through individual openings. However, only that tape selected by the tape selector is fed into the die and subsequently sheared by one punch stroke; see Fig. 4. The tape is punched from

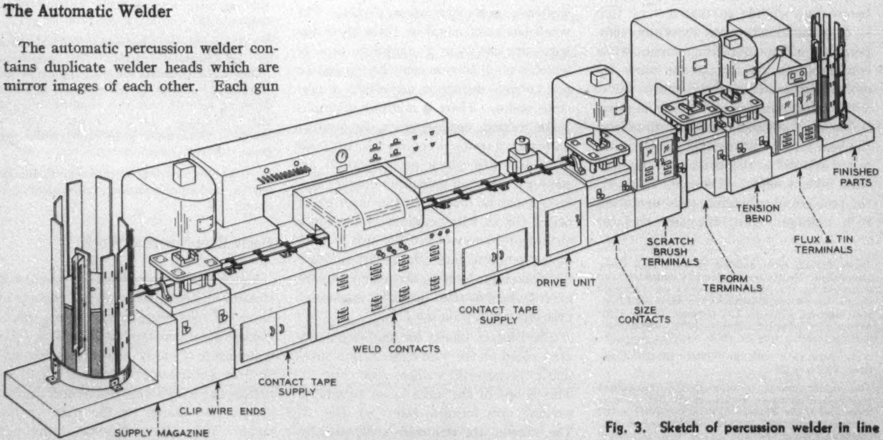

FINISHED PARTS
TENSION BEND
FLUX & TIN TERMINALS
SCRATCH BRUSH TERMINALS
FORM TERMINALS
DRIVE UNIT
SIZE CONTACTS
CONTACT TAPE SUPPLY
WELD CONTACTS
CONTACT TAPE SUPPLY
CLIP WIRE ENDS
SUPPLY MAGAZINE

Fig. 3. Sketch of percussion welder in line

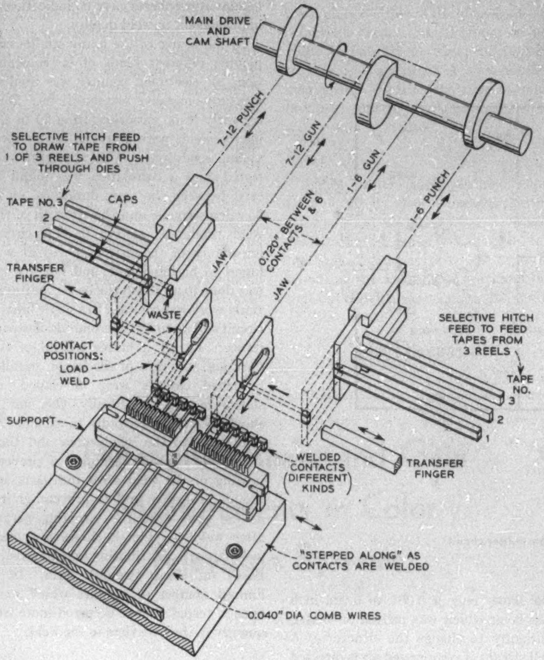

MAIN DRIVE AND CAM SHAFT
7-12 PUNCH
7-12 GUN
1-6 GUN
1-6 PUNCH
SELECTIVE HITCH FEED TO DRAW TAPE FROM 1 OF 3 REELS AND PUSH THROUGH DIES
TAPE NO. 3
2
1
CAPS
0.720" BETWEEN CONTACTS 1 & 6
JAW
JAW
TRANSFER FINGER
WASTE
CONTACT POSITIONS:
LOAD
WELD
SELECTIVE HITCH FEED TO FEED TAPES FROM 3 REELS
TAPE NO.
3
2
1
SUPPORT
WELDED CONTACTS (DIFFERENT KINDS)
TRANSFER FINGER
"STEPPED ALONG" AS CONTACTS ARE WELDED
0.040" DIA COMB WIRES

Fig. 4. Schematic of percussion welder

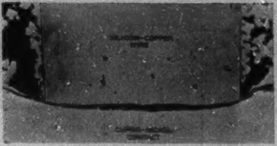

SILICON-COPPER WIRE

Fig. 5. Section through a weld

such a direction that the base metal is not dragged over the boundary line into the palladium zone. This avoids contamination of the palladium. As the contact is blanked out, the walls of a notch in the punch confine it on the precious metal sides to prevent distortion. The punch delivers the contact to a transfer finger at the end of the shearing stroke. This finger removes the contact from the punch notch and slides it through a guide channel into the waiting gun jaws.

WELDING GUNS

The welding gun is a light reciprocating member which carries two opposing steel fingers or jaws to receive the contact block from the transfer finger mentioned in the foregoing. The jaw opening is a few thousandths of an inch less than the nominal height of the contact; however, the edges are beveled so that, when a contact is pressed against them, they spring open and the contact enters and is held securely in place. After welding, the jaws are pulled off the contact. At the extreme return travel of the gun, any contact which might have remained in the jaws because it was not welded is removed by an ejector blade.

When in the loading position, a portion of the blade stops the travel of the contact through the jaw opening so that it is held in a uniform position and will be located on the wire with precision.

GUN MASS CONSIDERATION

During weld, the comb must be supported accurately to meet the close-contact location requirements. It must be supported securely to withstand the impact of the guns or weak welds may result. The mass of the gun is important. Evidence indicates that more uniform and higher weld strengths are obtained with lightweight guns. A magnesium gun weighing about 60 grams provided better welds than did the original steel gun weighing about 130 grams. A steel gun weighing about 30 grams is even more satisfactory. The indications are that the striking force should be as great as the part can stand. The large mass usually has a slower rebound time which must exceed the freezing time of the metal in the joint or it will be broken before it is cool enough to resist the stress. A striking force of approximately 75 pounds tends to loosen the wires in the plastic and to produce weak welds. The 60- or 30-gram guns propelled at about 40 inches per second at impact produce less than this force. The velocity during the arcing period is important for controlling the amount of heating. The freezing time is limited to the time the contact and wire end are forced together, or 1/2 cycle of vibration after impact.

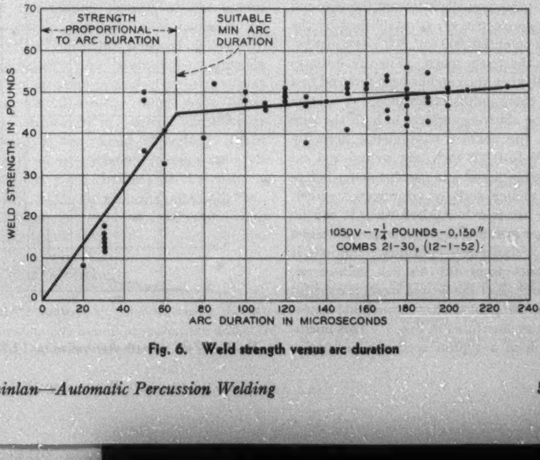


Fig. 6. Weld strength versus arc duration

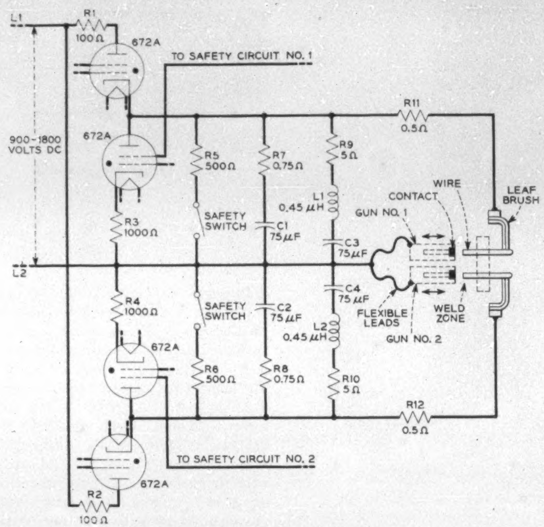


Fig. 7. Schematic of percussion welder circuit

Electrical Functions

Fundamentally, each weld circuit includes a capacitor which is charged during a small portion of each cycle and subsequently is discharged through a resistance in series with the weld. During the charging period, the contact and the wire end of the comb are separated electrically at the weld point. A multiple leaf brush under considerable spring force connects the one side of the circuit to the remote end of the individual wire to be welded. The gun, and through it the contact block, is connected to the other side of the circuit. After the cam frees the gun, the spring propells it toward the wire end and an electric arc is established by the high potential, 900 to 1,800 volts, just before the parts touch. The arc initiates itself when the gap has been reduced to a few thousandths of an inch. A portion of the abutting surfaces of both the wire and the contact base metal, normally 0.005 to 0.008 inch, are melted and expelled in liquid and gaseous states before the molten surfaces are forced together. The arc is extinguished when it can no longer melt and expel metal to maintain a gap. Under good operating conditions, it persists from 0.1 to 0.4 millisecond. Nearly all of the heated metal is expelled from the joint during the welding operation as illustrated in Fig. 5. This micrograph of a typical sectioned percussion weld shows only a 0.001 to 0.002 inch thick layer which was melted or heated sufficiently to change the structure. A small stream of compressed air is directed into the weld area to remove arc products so they will not interfere with the initiation of the arc during the next weld cycle. Limited tests made with nitrogen and helium atmospheres gave no indication of improvement in weld quality.

In the course of the many studies employing different kinds of instrumentation, an electronic counter was used to measure successive arc durations. During early tests, variations from 20 to 230 microseconds were observed for the prevailing conditions. Weld strengths were tested, and a correlation was found to exist between short arc durations, up to 65 microseconds, and the strength of the welds produced; see Fig. 6. The circuit impedance of the discharge circuit was found to be important, and much work was done to establish the best conditions. Early work was confined to a single branch circuit and up to 250 microfarads of capacitance. To get a better control of the duration and rate of discharge, parallel branched circuits were developed, as shown in Fig. 7. Ideally, the energy should be released rapidly to expand the weld heat to the whole area and then should be cut off abruptly to prevent heating of the weld after the parts are brought together and to prevent arcing at the jaws during a vibrating period after welding. Fig. 8 shows a typical weld strength distribution for this circuit, based on shear strength data. In a limited number of samples which were tension tested, failure occurred more frequently in the wire than in the weld.

Jaw Life

Jaw life is dependent upon many factors, but an important one is the prevention of accidental conditions which estab-

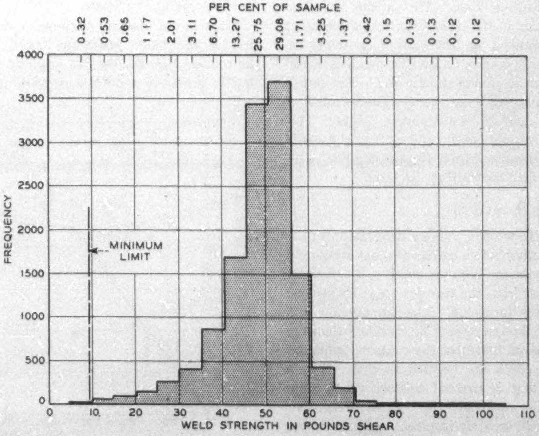


Fig. 8. Weld strength distribution for 13,200 samples tested from 600,000 production welds

lish an arc directly to the jaws. Each set of jaws is adjusted so that, under normal conditions, it will not touch the wire end and thereby discharge the capacitor if no contact is in the jaws. However, if only a very short length of contact metal is cut off and put in the jaws, it may start the arc but there may not be sufficient material or the material may not be held securely enough to keep the arc from burning the jaw surfaces. Another trouble condition is encountered when the end of the wire is misplaced so it touches a jaw. A safety circuit is provided with microswitches which trigger thyratrons to discharge the capacitors before the weld takes place if either of the aforementioned conditions occur. Signal lamps are provided to indicate at which microswitch trouble is occurring. A cam-actuated microswitch controls the time of safety test approximately 1/4 second before the weld would take place.

Safety

Safety for personnel from high voltage is provided by door switches, solenoid-released short-circuiting bars, and bleeder resistors on the capacitors. Safety from mechanical jams is provided by a slip clutch on the main drive gear, and by a pull-out clutch and automatic stop switch located in the piece-part transfer drive mechanism.

Conclusions

After making millions of welds by automatic percussion welding, it is found to be a method well suited to this job. Accuracy of location and good weld strength are obtained. This welding method is especially useful where speed and precision are desired and where joints must be made between dissimilar metals. Metals of high-heat conductivity and high-electric conductivity join readily by this method.

No Discussion

Differential Phase and Gain Measurements in Color Television Systems

H. P. KELLY
NONMEMBER AIEE

COLOR television has imposed new requirements on transmission circuits. The nature of the color signal is such that circuits used for its transmission must meet very stringent linearity requirements. The minimizing of differential phase and gain at the color carrier frequency is one of the more important of these new requirements.

Definitions

The National Television Systems Committee (NTSC) color television system is based on the principle that color may be adequately defined in terms of three characteristics: luminance, hue, and saturation. Luminance defines intensity or brightness and is the basis on which the present monochrome or black-and-white system of television operates. Hue defines the color in terms of whether it is red, blue, green, yellow, or the like. Saturation defines the degree to which the hue is mixed with white. For example, pink is a low saturation red. A high saturation red would be a brilliant crimson.

The color signal, therefore, must contain information as to these three characteristics. The color system uses the same type of signal to transmit luminance information as is used in the black-and-white system; therefore the transmission of this portion of the signal presents no new problems. The other two parts of the signal containing the hue and saturation information are added to the luminance information and comprise the new elements in the television signal. It is the necessity of transmitting three pieces of information instead of one, simultaneously, and without interaction or distortion, which has imposed the new requirements on transmission circuits. This situation is analogous to the transmission of two or more voice signals simultaneously in a carrier telephone circuit. If the circuit is perfectly linear, there is no difficulty in separating the various voice channels at the receiving end. If the circuit is not linear, the channels interact with one another and crosstalk occurs.

Color Information

In the case of color television the saturation and hue information are added to the luminance information in the form of a new signal called the color carrier. The amplitude of this signal represents the saturation of the color. A large amplitude represents high saturation or a brilliant color. Distortion of color saturation will occur if the gain of the transmission system to the color carrier is different for different amplitudes of the luminance signal. This variation in the amplitude transmission of the color carrier signal caused by variation in the amplitude of the luminance signal is called differential gain. The presence of differential gain in a system used to transmit color television may result in a picture in which some colors may appear dim or washed out while others may appear oversaturated.

The time or phase relationship of the color carrier signal to a reference synchronizing signal (color burst) determines the hue of the color. Distortions of hue will occur if the transmission system produces a different phase shift in the color carrier at different amplitudes of the luminance signal. This variation in phase shift caused by variations in amplitude of the luminance signal is called differential phase. The presence of differential phase in a system used to transmit color television results in a change of the hue of the colors.

The term "differential gain" was first proposed by S. Doba in a memorandum to the Video Techniques Committee of

---

Paper 54-334, recommended by the AIEE Television and Aural Broadcasting Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11-15, 1954. Manuscript submitted April 1, 1954; made available for printing July 9, 1954.

H. P. Kelly is with Bell Telephone Laboratories, Inc., Murray Hill, N. J.

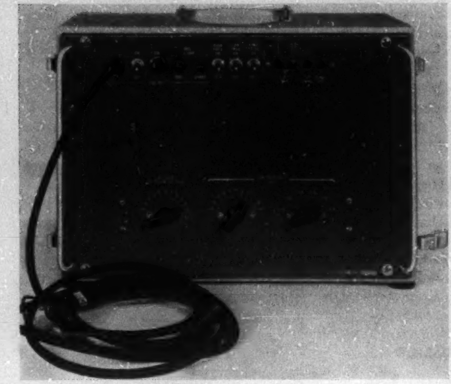


Fig. 1. Panel view transmitter

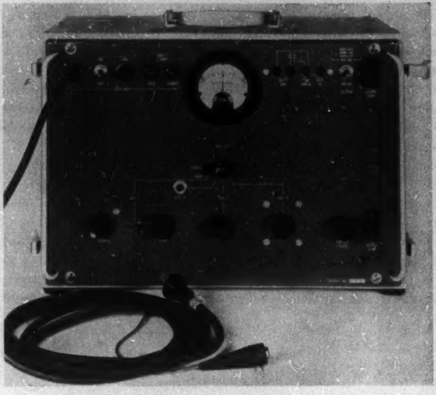


Fig. 2. Panel view receiver

the Institute of Radio Engineers. The method of measurement here presented follows along lines suggested by Doba. This method is one of several (e.g., Morrison's method,[1] and an alternate method used by R. C. Edson of the Bell Telephone Laboratories) which might be used to measure differential phase and gain, and was selected as giving the desired results quickly and accurately.

Equipment

The Bell Telephone Laboratories have developed a test set for measuring the presence of differential phase or gain at the color carrier frequency in television systems. It consists of two pieces of portable equipment: a transmitter and a receiver. The purpose of the transmitter is to generate a signal which simulates a color television signal. This signal is connected into the input of the system being tested. The purpose of the receiver is to examine the signal at the output of the system being tested, to determine if it has been subjected to differential phase or gain. The transmitter and receiver are physically independent of each other. The only connection between them is the circuit being tested. This independence is necessary for Bell System use since the transmitter and receiver might be separated by several thousands of miles of transmission networks.

Figs. 1 and 2 show panel views of the transmitting and receiving units. Fig. 3 shows a simplified block diagram of the system. The transmitter consists of three major parts: a luminance or black-and-white signal generator, a sync pulse generator and a constant color carrier generator. The luminance signal generator is a 15.750 kc sine-wave oscillator. The amplitude of this signal can be adjusted to represent a full range of luminance from dark to bright. The sync pulses are derived from the negative peaks of the luminance signal and therefore their positioning on the negative peaks is automatic. The color carrier generator is a 3.579545-megacycle crystal oscillator. Its output represents a constant hue and saturation signal. The three signals are combined at the output of the transmitter. The transmitter therefore sends a signal into the system under test which consists of a constant color carrier signal superimposed on a luminance signal which is continuously varied over the entire range from dark to bright. It can be seen that the effect of the luminance signal is to change the position of the color carrier in the amplitude range of the transmission system. It is this change in position which gives rise to differential phase and gain.

At the receiver the luminance signal and color carrier are separated by means of a filter. The luminance signal is applied to a cathode-ray oscilloscope in such a manner that it produces a left-to-right deflection of the beam. The left-to-right position of the beam, therefore, is an indication of the instantaneous amplitude of the luminance signal with the extreme left indicating the dark level and the extreme right indicating the bright level. The received color carrier has been drawn to illustrate the presence of differential gain. The color carrier signal containing any such variation which might have been produced in the system under test is split into two paths. The signal in one of these paths is transmitted through a sharp filter which removes any variation present and provides at the output of the filter an undistorted color carrier signal. This signal is used as a reference against which the distorted signal in the other path is compared. The comparison is made in the detector. The function of the detector can be switched so that a comparison may be made with respect to either amplitude or phase.

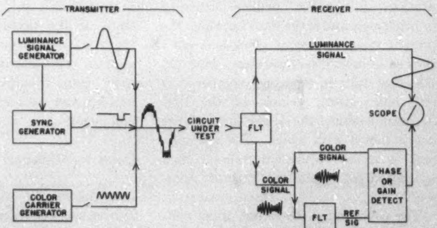


Fig. 3. System block schematic

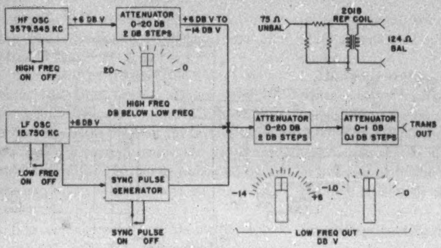

HF OSC 3579.545 KC
+6 DB V
ATTENUATOR 0-20 DB 2 DB STEPS
+6 DB V TO -14 DB V
HIGH FREQ ON OFF
20
0
HIGH FREQ DB BELOW LOW FREQ
LF OSC 15.750 KC
+6 DB V
LOW FREQ ON OFF
SYNC PULSE GENERATOR
SYNC PULSE ON OFF
75 Ω UNBAL
2018 REP COIL
124 Ω BAL
ATTENUATOR 0-20 DB 2 DB STEPS
ATTENUATOR 0-1 DB 0.1 DB STEPS
TRANS OUT
-14
+6
-1.0
0
LOW FREQ OUT DB V

Fig. 4. Transmitter block schematic

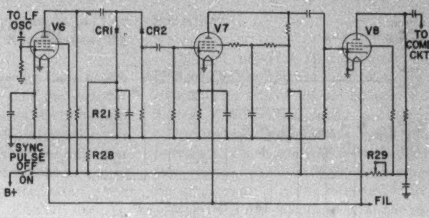

TO LF OSC
V6
CR1
CR2
V7
V8
TO COMB CKT
R21
SYNC PULSE OFF
ON
B+
R28
R29
FIL

Fig. 6. Sync generator

Measurements

When differential gain is being measured, the comparison is made on an amplitude basis. A voltage is obtained at the output of the detector, which is proportional to the difference in amplitude between the distorted and reference signals. In a similar manner, when differential phase is being measured, a voltage is obtained at the output which is proportional to the instantaneous difference in phase between the distorted and reference signals. In either case the output of the detector is applied to the oscilloscope in such a manner that it produces a up-and-down deflection of the beam. The amount by which the beam is deflected vertically is a measure of the amount of differential phase or gain present in the circuit under tests. The presentation of the complete measurement is therefore a line on a cathode-ray tube. This might be considered as a graph in which the amplitude of the luminance signal is plotted as a horizontal co-ordinate against differential gain or phase in the vertical co-ordinate. The sensitivity of the measurement is such that the amount of distortion which is considered just perceptible causes the beam to be deflected vertically 1 inch. This sensitivity enables us to measure, and hence correct, for distortions which are quite a bit below perceptibility. This is necessary because a long television network might consist of a number of shorter links. If the over-all network is to produce only a moderate or imperceptible degradation, then each link must be adjusted to be almost perfect in itself.

Fig. 4 is a block schematic of the transmitter showing the arrangement of the luminance generator, the sync generator, the color carrier generator, and the level-adjusting attenuators.

Luminance Signal Generator

A simplified schematic of the luminance or low-frequency signal generator is shown in Fig. 5. *V4*, *T3*, and *C22* comprise a 15.750-kc oscillator circuit. *V5* and associated components form a peak-to-peak detector circuit which is used to provide automatic volume control for the oscillator. The detector circuit is back-biased by a positive d-c reference voltage on cathode pin 8. So long as the peak-to-peak value of the 15.750-kc sine wave on the plate, pin 5 of *V4* does not exceed the d-c reference voltage on the detector, no d-c output is obtained at pin 4 of detector *V5*. However, when the peak-to-peak value does exceed the reference voltage, a negative d-c voltage is obtained at pin 4 of *V5* which is proportional to the amount of this excess. This voltage is fed back as bias to the grid of the oscillator tube and thus tends to limit the output of the oscillator. Potentiometer *R36* provides a means of varying the reference voltage to the detector and thereby the output of the oscillator. The reference voltage is supplied from a voltage reference gas tube. Changes in output of less than ±0.1 decibel (db) are obtained for ±10-per-cent variations in supply voltages. Frequency accuracy is of the order of ±1 per cent. A voltage which represents a small portion of the positive peak of the sine wave is obtained at pin 8 of the diode. This voltage is transmitted to the sync generator.

Sync Generator

A simplified schematic of the sync generator is shown in Fig. 6. This circuit consists of tubes *V6*, *V7*, and *V8*, *CR1* and *CR2* and associated components. The sine wave tips from the low-frequency generator are amplified and shaped into sync pulses in this circuit. *R29* provides a means of adjusting the amplitude of the sync pulse output.

Color Carrier Generator

A simplified schematic of the color carrier generator is shown in Fig. 7. It con-

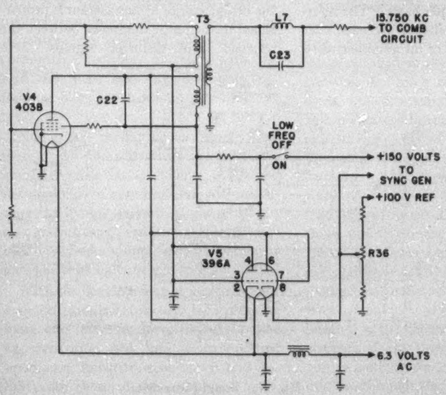
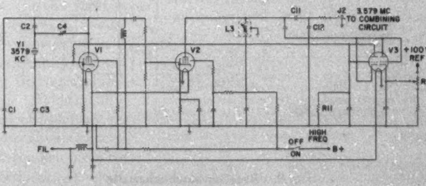

T3
L7
C23
15.750 KC TO COMB CIRCUIT
V4 403B
C22
LOW FREQ OFF
ON
+150 VOLTS
TO SYNC GEN
+100 V REF
R36
V5 396A
4
6
3
7
2
8
6.3 VOLTS AC

Fig. 5 (left). Low-frequency oscillator

Fig. 7 (below). High - frequency oscillator


C11
J2 3.579 MC TO COMBINING CIRCUIT
C2
C4
Y1 3579 KC
V1
L3
V2
C12
V3
+100V REF
R13
C1
C3
R11
HIGH FREQ
OFF
ON
B+
FIL

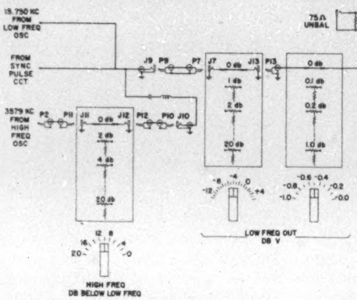

15,750 KC FROM LOW FREQ OSC
FROM SYNC PULSE CCT
3579 KC FROM HIGH FREQ OSC
T2
75 Ω UNBAL
124 Ω BAL
TRANS OUT
LOW FREQ OUT DB V
HIGH FREQ DB BELOW LOW FREQ

Fig. 8. Combining and output circuit

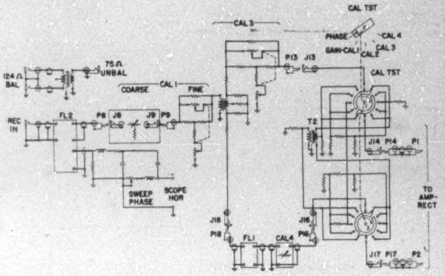

CAL TST
CAL 3
PHASE
GAIN-CAL 1
CAL 4
CAL 2
124 Ω BAL
75 Ω UNBAL
COARSE
CAL 1
FINE
REC IN
FL2
SWEEP PHASE
SCOPE HOR
T2
FL1
CAL4
TO AMP-RECT

Fig. 10. Input circuit

sists of a 3.579545 megacycle crystal oscillator *V1*, a buffer stage *V2*, and an automatic volume control circuit *V3*, which is similar in its operation to that described for the low-frequency oscillator. Potentiometer *R13* provides for adjusting the output of the oscillator.

The luminance, sync, and color carrier signals are combined as shown in Fig. 8. The output of the color carrier generator is transmitted through an attenuator before it is combined with the other two signals. This attenuator is designated "high freq—db below low freq" and is used to establish the ratio between the high- and low-frequency amplitudes. The combined output is transmitted through a pair of attenuators, designated "low freq out—dbv." These attenuators are calibrated in terms of the low-frequency or luminance signal output and cover a range of +6 decibels from 1 volt peak-to-peak (dbv) to −15 dbv in 0.1-db steps. A repeat coil is provided so that either a balanced 124-ohm output or an unbalanced 75-ohm output can be obtained.

Receiver

A block schematic of the receiver is shown in Fig. 9. The filter at the input consists of complementary low- and high-pass filters with their inputs connected in parallel. This filter is used to separate the luminance signal and the color carrier signal. The low-frequency luminance signal is obtained at the output of the low-pass section. The high-frequency color carrier signal is obtained at the output of the high-pass section. The sweep phase control in the luminance signal path provides for the adjustment of the horizontal sweep so that the extreme left-hand end of the oscilloscope trace represents the black level of luminance. The high-frequency color signal is split into two paths in the first hybrid coil.[2] The sharp crystal filter in one of the paths removes any distortion which might have been produced in the system under test and provides at its output an undistorted reference signal. This signal is used as a reference against which the distorted signal in the other path is compared. The "cal tst" switch provides the various circuit arrangements necessary for measuring differential phase or gain, or for calibrating adjustment. When measuring differential gain the reference and distorted signals are compared on an amplitude basis. The reference and distorted signals are each connected to the input of an amplifier-rectifier. The rectified outputs are compared in a bridge circuit and a voltage representing their difference is connected to the vertical input of an oscilloscope. The two amplifier-rectifiers are adjusted to have equal sensitivity by means of the *cal 2* control. Therefore, if at any instant the reference and distorted signals are equal in amplitude, the output from the bridge circuit is zero.

Phase Measurements

When differential phase is being measured the reference and the distorted signals are compared on a phase basis. The two signals are connected into a pair of conjugate branches of a hybrid coil. The two outputs from the other pair of conjugate branches represent the phasor sum and the phasor difference of the two signals. The phasor sum is connected to the input of one of the amplifier-rectifiers and the phasor difference to the other. The outputs are combined in the same manner as for differential gain. The output from the bridge circuit at any instant is proportional to the phase difference between the reference and distorted signals. Zero

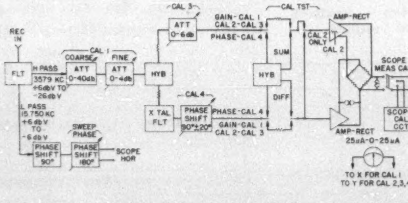

REC IN
FLT
H PASS 3579 KC +6dbV TO -26dbV
L PASS 15,750 KC +6 dbV TO -6 dbV
CAL 1
COARSE
FINE
ATT 0-40db
ATT 0-4db
HYB
X TAL FLT
CAL 3
ATT 0-6db
CAL 4
PHASE SHIFT 90°±20°
GAIN-CAL 1 CAL 2-CAL 3
PHASE-CAL 4
CAL TST
HYB
SUM
DIFF
CAL 2 ONLY
CAL 2
AMP-RECT
AMP-RECT 25μA-0-25μA
SCOPE MEAS CAL
SCOPE VERT
SCOPE CAL CCT
SWEEP PHASE
PHASE SHIFT 90°
PHASE SHIFT 180°
SCOPE HOR
TO X FOR CAL I
TO Y FOR CAL 2,3,4

Fig. 9. Receiver block schematic

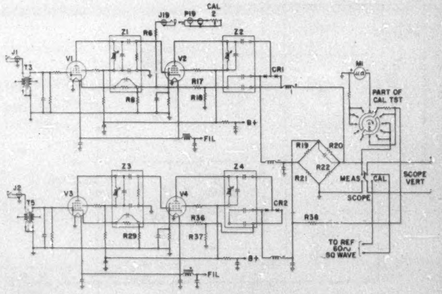

J19
P19
CAL 2
Z1
R6
Z2
J1
T3
V1
V2
R17
R18
R8
CR1
M1
PART OF CAL TST
B+
FIL
R19
R20
R21
R22
MEAS
CAL
SCOPE
SCOPE VERT
Z3
Z4
J2
T5
V3
V4
R36
R29
R37
CR2
R38
TO REF 60~ SQ WAVE
B+
FIL

Fig. 11. Amplifier-rectifier

output is obtained from the bridge when the two signals are 90 degrees out of phase.

The *cal 1* position of the switch provides the same circuit arrangement as for gain. The *cal 1* control is used to adjust the input level for the proper value of rectified current from the reference signal amplifier-rectifier as indicated by a reading of 20 microamperes on the panel meter.

The *cal 2* position of the switch provides for connecting the reference signal to the input of both amplifier-rectifiers. The panel meter is connected to the output of the bridge circuit. The adjustment of the *cal 2* control for zero output from the bridge then provides equal sensitivity in the two amplifier-rectifiers.

The *cal 3* position provides the same circuit arrangement as for gain except that the meter is connected to the output of the bridge. The *cal 3* control is then adjusted for equal transmission in the two paths as indicated by a zero-reading of the meter.

The *cal 4* position is the same as the phase position except that the meter is connected to the output of the bridge circuit. The *cal 4* control is adjusted in this position to obtain a 90-degree difference in phase shift in the two paths as indicated by a zero-reading of the meter.

Amplifier-Rectifiers

Simplified schematics of the input circuits and amplifier-rectifiers are shown on Figs. 10 and 11. Each of the two amplifier-rectifier circuits consists of a 2-stage 3,579-kc amplifier terminated in a varistor rectifier. Each circuit has approximately 26 db of negative feedback. Since the rectifier circuit is included in the feedback loop, a high degree of linearity and stability is obtained.

Sensitivity

This set provides approximately 50 millivolts to the vertical input of the oscilloscope for 1 db of differential gain or 5 degrees of differential phase. A number of oscilloscopes are available which will provide a 2-inch vertical deflection for this voltage. A 50-millivolt calibrating signal is available in the receiver for adjusting the vertical sensitivity of the oscilloscope.

References

1. TEST INSTRUMENTS FOR COLOR TELEVISION, **W. C. Morrison, K. Karstad, W. L. Behrend.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., Jan. 1954, vol. 42, pp. 247–58.

2. COMMUNICATION AND ELECTRONICS (ELECTRICAL ENGINEERS HANDBOOK), **Harold Pender, Knox McIlwain.** John Wiley & Sons, Inc., New York, N. Y., 1950.

◆

No Discussion

Derivative-Controlled Magnetic Amplifiers

A. D. SCHNITZLER
AFFILIATE MEMBER AIEE

ALTHOUGH magnetic amplifiers have been in use for a great many years, it is only in the past few that they have widespread attention. Their development has been vastly accelerated by the development of high quality core materials and rectifiers, and the recognition of the vital role that rectifiers may play in magnetic-amplifier circuits. Thus the self-saturating magnetic amplifier[1,2] incorporating rectifiers in series with the load windings, was invented and was found to have a high gain to time constant ratio.

Numerous modifications have been made upon the self-saturating magnetic amplifier since its invention. After the development of core material with a rectangular hysteresis loop, Ramey,[3] by incorporating rectifiers in the control circuit, invented an amplifier with high-gain and inherent half-cycle response.

With the development of more complex magnetic-amplifier circuits incorporating more and more rectifiers, it has become apparent that new limitations imposed upon gain, linearity, sensitivity, and stability are those of the component rectifiers. For example, the incorporation of rectifiers in the control circuit of a half-cycle response amplifier sets a lower limit upon the signal voltage of about 1 volt.

Recently, this limitation has been removed by Scorgie,[4] who invented a magnetic-amplifier circuit which incorporates a combination of positive and negative feedback in such a way as to obtain high-gain and half-cycle response without the use of rectifiers in the control circuit. In addition, by using core material with a large value of maximum permeability μ_m, he has obtained linear amplification of signals on the order of 10^{-3} volts while at the same time preserving the half-cycle response. The dependence of the forward resistance of the rectifiers in the load circuit upon the current through them and the largest value of μ_m available set this lower limit upon signal voltage. However, for small values of signal voltage, self-saturating magnetic amplifiers are subject to drift and instability, because of the dependence of the back resistance of the rectifiers upon temperature, which in turn is dependent upon the current through the rectifiers.

This paper describes a high-gain fast-response drift-free magnetic amplifier for operation with such low-level low-impedance signal sources as thermocouples. This amplifier is, however, one of a class to which the rule applies that the input impedance increases as the square of the number of turns of the control winding. The breadboard model described in this paper was found to have a response time of 1 to 2 cycles of the power supply frequency, a constant voltage gain of 700, a power gain ranging from 10,000 to 105,000 and to be capable of operating from a signal source with up to 5 ohms of impedance without affecting the fast response time. Freedom from drift and high gain are obtained by use of a combination of positive and negative feedback in such a way that the first stage of the 2-stage amplifier operates on the average increment of signal voltage over a half-cycle, while the second stage amplifies and sums the increment already amplified by the first stage. In a 2-stage amplifier, drift is almost entirely due to variations in the parameters of the first stage. Here, the first stage functions as a derivative amplifier and the second stage as an integrator. Since the input stage operates upon the average increment of

---

Paper 54-375, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted March 30, 1954; made available for printing August 13, 1954.

A. D. SCHNITZLER is with the Naval Research Laboratory, Washington, D. C.

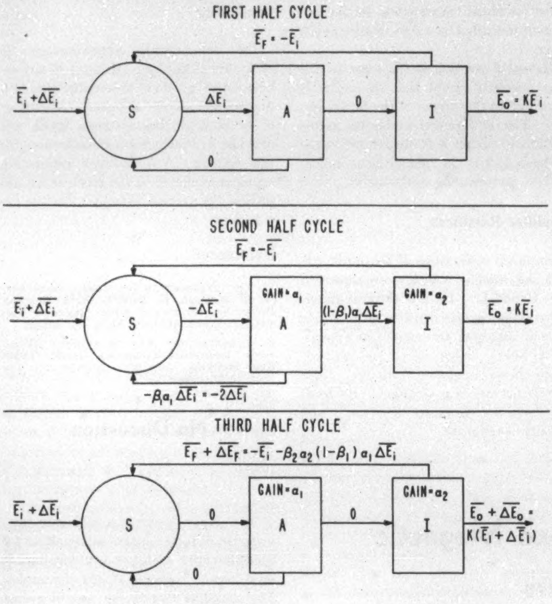

FIRST HALF CYCLE
$\overline{E_F} = -\overline{E_i}$
$\overline{E_i} + \overline{\Delta E_i}$
S
$\overline{\Delta E_i}$
A
0
I
$\overline{E_o} = K\overline{E_i}$
0
SECOND HALF CYCLE
$\overline{E_F} = -\overline{E_i}$
$\overline{E_i} + \overline{\Delta E_i}$
S
$-\overline{\Delta E_i}$
GAIN = α_1
A
$(1-\beta_1)\alpha_1\overline{\Delta E_i}$
GAIN = α_2
I
$\overline{E_o} = K\overline{E_i}$
$-\beta_1\alpha_1\overline{\Delta E_i} = -2\overline{\Delta E_i}$
THIRD HALF CYCLE
$\overline{E_F} + \overline{\Delta E_F} = -\overline{E_i} - \beta_2\alpha_2(1-\beta_1)\alpha_1\overline{\Delta E_i}$
$\overline{E_i} + \overline{\Delta E_i}$
S
0
GAIN = α_1
A
0
GAIN = α_2
I
$\overline{E_o} + \overline{\Delta E_o} = K(\overline{E_i} + \overline{\Delta E_i})$
0
$K = \alpha_2(\alpha_1 - 2) - 1$

Fig. 1. Block diagram

signal voltage per half-cycle, rather than upon the entire signal voltage as in the case of a conventional 2-stage amplifier, the effect of variations in the back resistance of the rectifiers upon the output is diminished by the ratio of the increment per half-cycle to the entire signal voltage. Drifting by, the integrator is degenerated by 100 per cent delayed negative feedback. By delaying the negative feedback through the action of a filter or a reactive load, the output voltage of the amplifier can be made proportional to the sum of the signal voltage and its first derivative. This effectively reduces the total time constant of a system incorporating the amplifier.

General Theory

A block diagram depicting the functional relations between average voltages during successive half-cycles of the power supply voltage is shown in Fig. 1. The three separate diagrams indicate the voltages during three successive half-cycles following the initiation of a step increase in signal voltage from $\overline{E}_i$ to $\overline{E}_i + \overline{\Delta E}_i$. The bar over the voltage symbol indicates the average voltage over a half-cycle.

The block labeled S consists of a resistor network for summing the three input voltages. These voltages are the signal voltage and two negative feedback voltages. It will be shown later that the sum of these three voltages, which constitutes the input to block A, is proportional to the derivative of the signal voltage. Block A may consist of either two full-wave or two half-wave magnetic amplifiers connected in a push-pull arrangement so as to achieve amplification of both positive and negative increments of signal voltage.

The block labeled I is a voltage integrator, i.e., a conventional self-saturating magnetic amplifier with sufficient external positive feedback to compensate for the power losses which are due to rectifier leakage and core losses.

A general understanding of the operation of the amplifier can be obtained by referring to the block diagram. It is assumed that enough time has elapsed after the termination of any previous transients for the output of the amplifier to have reached a steady state. It follows from this assumption that the input to the integrator (but not the output), and the output and input of amplifier A must be zero. Therefore, the feedback voltage from the integrator to the resistor summing network S must cancel the signal voltage $\overline{E}_i$.

It is now assumed that at the beginning of a half-cycle of the supply voltage an increment $\overline{\Delta E}_i$ of signal voltages appears at the input to the summing network S. Due to the addition of successive half-cycle delays inherent in the operation of amplifier A and integrator I, the feedback voltage $\overline{E}_F$ remains unchanged during the first and second half-cycles after the appearance of the increment of signal voltage. Likewise, the feedback voltage from amplifier A remains zero during the first half-cycle after the appearance of the increment of signal voltage. Hence a net voltage $\overline{\Delta E}_i$ appears in the control circuit of amplifier A. This voltage determines the output of amplifier A during the second half-cycle.

Since the voltage gain of amplifier A is α_1 during the second half-cycle, its total output voltage is $\alpha_1\Delta E_i$. Of this total voltage a fraction β_1 appears across one of the resistors of the summing network S and the fraction $1-\beta_1$ appears at the input to the integrator. By selecting β_1 so that

$$-\beta_1\alpha_1\overline{\Delta E}_i = -2\overline{\Delta E}_i \qquad (1)$$

the input to the amplifier A is made zero during the second half-cycle and hence the output from the amplifier during the third half-cycle is zero. The factor 2 in equation 1 arises from the fact that the flux transferred from the gating core to the resetting core of A is diminished by an amount proportional to $-\overline{\Delta E}_i$, the input voltage during the first half-cycle. Therefore $-\Delta E$ must be provided to ensure zero output from A during the third half-cycle.

Then, during the third half-cycle, if the voltage gain of the integrator is α_2, the total output voltage is increased by $\alpha_2(1-\beta_1)\alpha_1\overline{\Delta E}_i$. Of this voltage increase, the fraction β_2 is caused to appear as additional negative feedback voltage to the resistor network S, and the fraction $1-\beta_1$ is caused to appear as an increase in load voltage. By selecting β_2 so that

$$-\beta_2\alpha_2(1-\beta_1)\alpha_1\overline{\Delta E}_i = -\overline{\Delta E}_i \qquad (2)$$

the input to amplifier A again is made zero during the third half-cycle. Since both the input and output voltages of amplifier A are zero during the third half-cycle, the output of the integrator cannot change further until at least 1 cycle after the initiation of a new increment of signal voltage. Thus a steady state has been established 1 cycle after the initiation of the incremental change in signal voltage.

Equations 1 and 2 equating the increment in feedback voltage to the increment of input voltage determine β_1 and β_2 as functions of α_1 and α_2. From the first equation

$$\beta_1=\frac{2}{\alpha_1} \tag{3}$$

Substituting this value of β_1 into equation 2 gives for β_2

$$\beta_2=\frac{1}{\alpha_2(\alpha_1-2)} \tag{4}$$

The increment of voltage across the load is given by

$$\overline{\Delta E_0}=(1-\beta_2)\alpha_2(1-\beta_1)\alpha_1\overline{\Delta E_i} \tag{5}$$

Substituting in equation 5 the expressions for β_1 and β_2, the voltage gain of the amplifier, defined as the ratio of $\overline{\Delta E_0}$ to $\overline{\Delta E_i}$ is

$$\frac{\overline{\Delta E_0}}{\overline{\Delta E_i}}=\alpha_2(\alpha_1-2)-1 \tag{6}$$

If α_1 and α_2 are large compared to unity, the total voltage gain of the amplifier is approximately $\alpha_1\alpha_2$.

In this description of the functional relationship between the blocks, it was assumed that enough time had elapsed after the termination of any previous transients for the output of the amplifier to have reached a steady state. This means that the signal voltage is assumed to have been constant for at least 3 half-cycles before the initiation of the new transient assumed in the analysis. Let T_{-2} designate the second half-cycle and T_{-1} the first half-cycle before initiation of a transient and T_n the nth half-cycle after initiation of the new transient. Then, in accordance with the initial conditions

$$\bar{E}_i(T_{-2})=\bar{E}_i(T_{-1}) \tag{7}$$

$$\bar{E}_i(T_{-1})+\overline{\Delta E_i}=\bar{E}_i(T_1) \tag{8}$$

Now suppose the input voltage continues to change so that its value during the second half-cycle in general differs from that during the first half-cycle. Then the input to amplifier A during the second half-cycle after the transient is

$$E_i(T_2)-E_i(T_{-1})-[E_i(T_1)-E_i(T_{-2})]$$
$$=E_i(T_2)-E_i(T_1) \tag{9}$$

The first term on the left of equation 9 is the signal during the second half-cycle. The second term is the feedback from the integrator delayed 1 cycle, and the third term is the feedback from amplifier A delayed 1 half-cycle.

Similarly, during the third half-cycle, the input to amplifier A is given by

$$\bar{E}_i(T_3)-\bar{E}_i(T_1)-[\bar{E}_i(T_2)-\bar{E}_i(T_1)]$$
$$=\bar{E}_i(T_3)-\bar{E}_i(T_2) \tag{10}$$

In general, then, the input to amplifier A during the nth half-cycle is given by $\bar{E}_i(T_n)-\bar{E}_i(T_{n-1})$. Since the time derivative is defined by

$$\frac{dE_i(t)}{dt}=\underset{t\rightarrow t_0}{\text{limit}}\frac{E_i(t)-E_i(t_0)}{t-t_0} \tag{11}$$

it is apparent that the input to amplifier A is approximately the time derivative of the signal voltage multiplied by the half period of the power supply voltage. Thus

$$\bar{E}_i(T_n)-\bar{E}_i(T_{n-1})=\frac{\bar{E}_i(T_n)-\bar{E}_i(T_{n-1})}{T_n-T_{n-1}}\times$$
$$(T_n-T_{n-1}) \tag{12}$$

$$=\frac{\bar{d}_{Ei}(T_n)}{dt}(T_n-T_{n-1}) \tag{13}$$

Since T_n-T_{n-1} is the half-period of the power supply voltage and a constant, the input to amplifier A is equal to the average value per half-cycle of the first derivative of the signal voltage. Because the change in signal voltage per half-cycle is really always less than the total signal voltage, the amplification requirements of amplifier A are correspondingly reduced. It will be shown also that the input power and the tendency for the amplifier to drift are reduced.

An even greater reduction of the amplification requirements of the amplifier can be achieved by adding successive stages in such a way that the input to the first stage contains successively higher order derivatives of the signal voltage. This can be seen best by considerng the Taylor series expansion of a function $E_i(t)$. Thus

$$E_i(t)=\sum_{K=0}^{n}\frac{1}{K!}(t-t_0)E_i^{(K)}(t_0)+R_n \tag{14}$$

where R_n is the error.

$$R_n=\sum_{K=n+1}^{\infty}\frac{1}{K!}(t-t_0)^K E_i^{(K)}(t_0) \quad 0\leq t-t_0<1 \tag{15}$$

If $n=0$

$$E_i(t)=E_i(t_0)+R_0 \tag{16}$$

The first stage of a 2-stage amplifier has the error term as its input. Since the error contains only first- and higher order derivatives, the input is zero if the signal is constant. If $n=1$

$$E_i(t)=E_i(t_0)+(t-t_0)\frac{dE(t_0)}{dt}+R_1 \tag{17}$$

In the case of a 3-stage amplifier, the error term contains only second- and higher order derivatives. If the first derivative of the signal is constant, the input to the first stage is zero.

In general for an $n+2$-stage amplifier, if the nth derivative of the signal is constant, the error term which is the input to the first stage is zero. If the signal and its derivatives are continuous and not zero, the error term approaches zero as more and more terms are included. Since the error term is the input to the first stage, the input requirements of the first stage can be reduced by adding more and more stages. However, an additional half-cycle delay is introduced by each additional stage. Also the number of cores and rectifiers required to make up the amplifier increases geometrically with the number of stages.

It would seem that the 2-stage amplifier with its high gain and stability and only 1-cycle delay should be useful in many applications. The 3-stage amplifier incorporating more cores and rectifiers than the 2-stage amplifier could be useful in those applications where extremely high gain and sensitivity are desired. Only the 2-stage amplifier is given a detailed description in this paper.

Description of Circuit

The 2-stage derivative-controlled magnetic amplifier utilizing full-wave derivative amplifiers is shown in Fig. 2. This circuit will now be described in detail. Later a simplified version embodying half-wave derivative amplifiers will be described.

In Fig. 2, cores I and II with associated circuitry are the positive derivative amplifier; cores III and IV, the negative derivative amplifier; and cores V and VI, the integrator. The negative feedback from the positive and negative derivative amplifiers appears across resistors R_1 and R_2 respectively. The negative feedback voltage from the integrator appears between terminals X_3 and X_4, which are across resistor R_9. The three blocks labeled I_1, I_2, and I_3 respectively, are constant current sources which provide bias currents to the cores. Rectifiers r_9 and r_{10} in conjunction with resistors R_3 and R_4 serve to isolate the load circuits of the two derivative amplifiers from each other. Resistor R_6 is part of the external positive feedback circuit of the integrator. Resistor R_7 is the load assumed to be purely resistive. Resistors R_8 and R_9 and the low-pass filter are part of the negative feedback circuit of the integrator. The low-pass filter serves to attenuate second- and higher-order harmonics of the power supply frequency in the negative feedback to the control circuit of the derivative amplifiers. The filter also delays the negative feedback voltage. That this leads to an improvement in sensi-

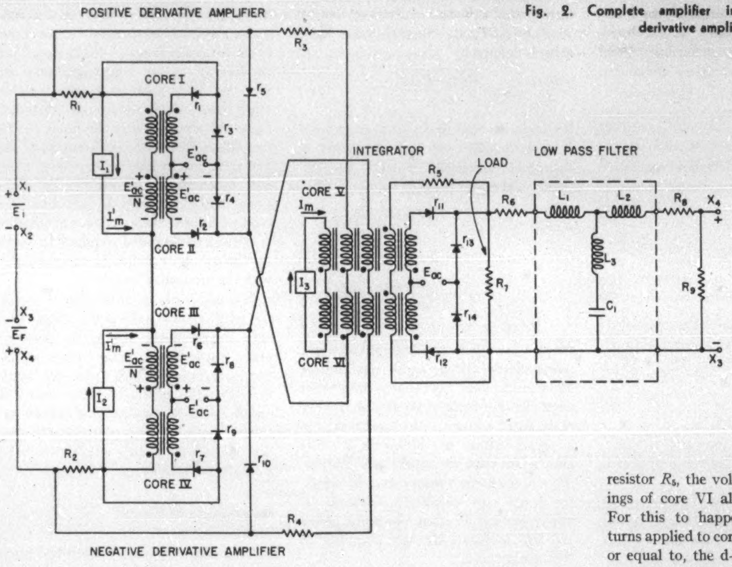


Fig. 2. Complete amplifier incorporating full-wave derivative amplifiers

tivity, stability, and time of response is shown after the operation of the derivative amplifiers is described.

The Integrator

The operation of the integrator will be considered first, followed by consideration of the operation of the derivative amplifiers. A simplified form of this circuit, shown in Fig. 3, is similar to that described by reference 2. The amplitude of the power supply voltage is just large enough to swing the flux in either core from knee to knee of the hysteresis loop. The core material is assumed to be of the rectangular hysteresis loop type. This assumption is later abandoned since the amplifier will operate using any material with a high ratio of remanent flux to coercive force. The constant current source I_3 supplies a steady direct current equal to the magnetizing current of the N_1 winding at the power supply frequency. The rectifiers are assumed to have a negligible forward resistance and a constant backward resistance. At first, it will be assumed that the input voltage $\overline{E}_d$ is zero. If the polarity of the supply voltage is that indicated in Fig. 3, core V is on its gating half-cycle and core VI on its resetting half-cycle.

Two modes of operation determined by the magnetic state of the gating core occur each half-cycle. During the first part of the half-cycle, the gating core is unsaturated, and inductive coupling exists between the windings of that core. During the latter part of the half-cycle, the gating core is saturated and the windings are effectively decoupled from one another.

The impedance of the load winding N_5 of the gating core while the core is unsaturated is much greater than the resistance of R_7'. Hence, before core V saturates, the voltage across the load winding on core V is e_{ac}, (i.e., the instantaneous value of the supply voltage). The following argument can be used to establish that the voltage across the load winding N_5 of core VI is also e_{ac}: First, it is to be noted that the current supplied by the constant current source is equal to the magnetizing current of the N_1 winding. Therefore, the sum of the ampere turns due to currents flowing in the other three windings on the core must be equal to zero. But the voltages across the windings of core V are such as to cause currents to flow into the undotted ends of all three windings of core VI. Hence the currents must be zero and the voltages across corresponding windings of cores V and VI must be of equal magnitude and of opposite polarity.

After core V saturates, the voltages across all its windings drop to zero and the supply voltage appears across the load resistor R_7'. By adjusting the feedback resistor R_5, the voltages across the windings of core VI also can be made zero. For this to happen, the total ampere turns applied to core VI must be less than, or equal to, the d-c magnetizing ampere turns of the core. Hence, by adjusting the feedback ampere turns to be greater than or equal to the rectifier (r_{12}), leakage ampere turns plus the difference between the magnetizing ampere turns at the power supply frequency and its value at zero frequency or direct current the voltages across the windings of core VI drop to zero while core V is saturated.

Since the voltages across similar windings of both cores are instantaneously equal, the flux change necessary to saturate core V must just equal the flux change in resetting core VI. It is safe to assume that the flux in the resetting core each half-cycle is initially at the upper knee of the hysteresis loop since the amplitude of e_{ac} is large enough to swing the flux from knee to knee of the hysteresis loop. If the flux in the gating core is at some arbitrary value at the beginning of the half-cycle, the flux in the resetting core will assume this same value by the end of the half-cycle. Thus, not only are the flux changes in the two cores equal, but the value of flux in one core at the beginning of the half-cycle becomes the value of flux in the other core by the end of the half-cycle. At the beginning of the next half-cycle, when the polarity of the supply voltage changes and thereby interchanges the roles of the gating core and the resetting core, the initial conditions of the two cores have also been interchanged. This means that the operation during each half-cycle is a repetition of that during the preceding half-cycle. Thus, the average output voltage over a half-cycle is constant and is given by

$$\bar{E}_0 = \frac{E_{ac}}{\pi}(1+\cos\theta) \tag{18}$$

where θ is the phase angle at which the gating core saturates. The change in flux in either core is given by

$$C_1\Delta\phi = \frac{E_{ac}}{\omega}(1-\cos\theta) \tag{19}$$

where C_1 is a constant depending upon the number of turns in the winding and units. Making use of this equation, the average output voltage can be written as

$$\bar{E}_0 = \frac{2E_{ac}}{\pi} - \frac{\omega}{\pi}C_1\Delta\phi \tag{20}$$

Now suppose the input voltage $\bar{E}_d$ is not zero, but has some magnitude, and the polarity shown in Fig. 3. How this voltage is divided by the two N_2 windings of the two cores and resistor R_3 depends upon the relative impedances of the three elements. The ratio of the impedance of the N_2 winding of core VI to the N_2 winding of core V is on the order of the ratio of the backward resistance of rectifier r_{12} to the resistance of R_7. In practice, this is a large number and therefore E_d is divided between R_3 and the N_2 winding of core VI. For maximum efficiency of power transfer from the derivative amplifiers to the integrator, R_3 is chosen so that approximately one-third of the input voltage is dropped across the N_2 winding of core VI.

The flux change in the resetting core is now equal to the flux change in the gating core minus a flux proportional to the time integral of E_d over a half-cycle. During the second half-cycle after initiation of E_d, since the gating core was the resetting core during the first half-cycle, the net flux change in the resetting core is decreased by twice as much as during the first half-cycle. For the nth half-cycle this can be written

$$\Delta\phi_n = \Delta\phi_1 - C_2\sum_{k=1}^{n-1}(\bar{E}_d)_K \tag{21}$$

where C_2 is a constant having the units of time. According to equation 20, during the nth half cycle

$$(\bar{E}_0)_n = 2\frac{E_{ac}}{\pi} - \frac{\omega}{\pi}C_1\Delta\phi_n \tag{22}$$

Substituting $\Delta\phi_n$ from equation 21

$$(\bar{E}_0)_n = 2\frac{E_{ac}}{\pi} - \frac{\omega}{\pi}C_1\Delta\phi_1 + \frac{\omega}{\pi}C_2C_1\sum_{k=1}^{n-1}(\bar{E}_d)_K \tag{23}$$

This can also be written

$$(\bar{E}_0)_n = (\bar{E}_0)_1 + C_3\sum_{k=1}^{n-1}(\bar{E}_d)_K \tag{24}$$

where C_3 is a new constant. Each half-cycle then $\bar{E}_0$ increases an amount proportional to the average value of E_d during the preceding half-cycle.

Likewise $\bar{E}_0$ decreases an amount each half-cycle proportional to $\bar{E}_d$ during the previous half-cycle when the polarity of E_d is the reverse of that shown in Fig. 3. Thus, the circuit normally operates as an integrator.

However, if the polarity of E_d is the reverse of that shown in Fig. 3, a linear relation between the output voltage and the integral of the input voltage can exist only if

$$E_d = 0 \quad 0 \leq t < T_s \tag{25}$$

$$E_d \leq e_{ac} \quad T_s \leq t \leq \frac{T}{2} \tag{26}$$

where T_s is the time at which the gating core saturates. If the voltage across the load winding of the resetting core exceeds e_{ac}, current flows through rectifier r_{12} in the forward direction. Thus, the impedance of the N_2 winding of the resetting core is greatly reduced and nearly all of the input voltage is dropped across R_3 before the gating core saturates and the input voltage in excess of e_{ac} is dropped across R_3 after the gating core saturates. The conditions on input voltage which insure linearity between the output voltage and the integral of the input voltage are met by proper design. This is possible since the input voltage wave form is from the output of a magnetic amplifier and hence is a pulse occurring during the latter part of the half-cycle.

Derivative Amplifiers

The full-wave derivative amplifiers are shown in Fig. 4. It has been found that in practice this circuit can be greatly simplified and a simplified circuit is presented later in this paper. However, the circuit shown in Fig. 4 will be used to elucidate the theory of operation.

The assumptions made here are similar to those made in describing the operation of the integrator. The power supply voltages E_{ac}' are just large enough to

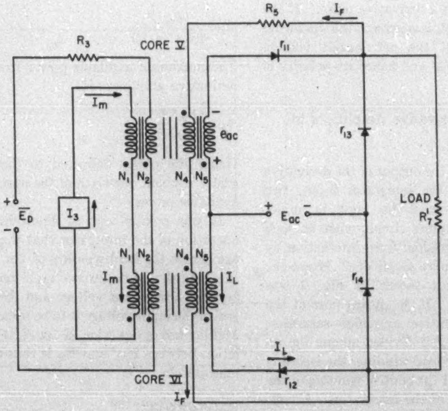


Fig. 3. Integrator

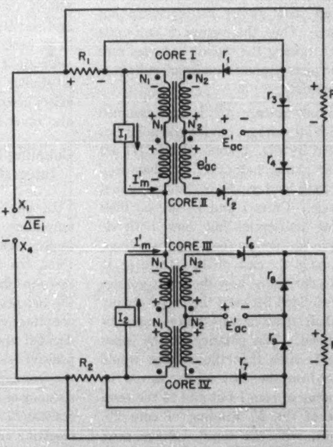


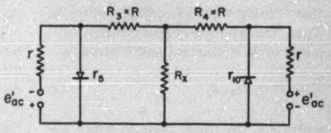
Fig. 4. Full-wave derivative amplifiers

swing the flux in any of the four magnetically similar cores from knee to knee of the hysteresis loop. For simplicity of analysis, the core material is assumed to be of the rectangular hysteresis loop type, an assumption which again, as in the case of the integrator, can be abandoned in practice. The constant current sources I_1 and I_2 supply steady currents equal to the magnetizing currents of the N_1 windings of the cores at the power supply frequency. The rectifiers are assumed to have negligible forward resistance and a constant backward resistance. First the input voltage $\overline{\Delta E_t}$ is assumed to be zero. When the polarities of the supply voltages are such as shown in Fig. 4, cores II and III are on their gating half-cycle while cores I and IV are on their resetting half-cycle.

Before the gating cores saturate, a voltage equal to approximately e_{ac}' exists across their load windings and the voltage across the load resistors is approximately zero. The voltage across the load windings of the resetting cores must also be e_{ac}' since the constant current sources supply the magnetizing currents of the N_1 windings on the cores. If the voltage were less than e_{ac}' leakage currents through rectifiers r_1 and r_7 would cause the ampere turns applied to the cores to exceed the ampere turns required to maintain the voltage at e_{ac}'. Thus a contradiction results. Similarly, a contradiction results if the voltage is assumed to exceed e_{ac}'. Since there is no positive feedback here, as there was in the case of the integrator, the flux goes to the lower knee of the hysteresis loop of each resetting core in a half-cycle. These same two cores on their gating half-cycle do not saturate. Thus the output 1 half-cycle after energizing the circuit becomes zero and remains zero until an input voltage is introduced.

An input voltage $\overline{\Delta E_t}$ is now assumed and with all voltages having the polarities indicated by Fig. 4, cores II and III are on their gating half-cycle and have e_{ac}' across their load windings during the first half-cycle. Cores I and IV are on their resetting half-cycle, but now both do not have e_{ac}' across their load windings. Again, as in the case of the integrator, in order to determine how the input voltage will be divided between the two N_1 windings, their relative impedances must be determined. The polarity of the input voltage is such that the voltage would subtract from the reset voltage across the N_1 winding of core I but add to the reset voltage of the N_1 winding of core IV. However, with the input voltage zero, N_1e_{ac}'/N_2 exists across each of the windings. The impedance in the load winding of core I, if the voltage is less than e_{ac}', is the backward resistance of r_1 while the impedance in the load winding of core IV, if the voltage is greater than e_{ac}', is the forward resistance of r_7. Hence, nearly all of the input voltage is dropped across the N_1 winding of core I. Thus the voltage across the load winding of core I is $\overline{E_{ac}}'-(N_2/N_1)\overline{\Delta E_t}$ during the first half-cycle, and only core IV will be fully reset.

Fig. 5. Equivalent circuit of interstage coupling



During the second half-cycle, when the gating cores are interchanged with the resetting cores, similar conditions hold but with different cores, until core I which is now on its gating half-cycle saturates. While core I is saturated, a negative feedback voltage exists across R_1. The ratio of R_1 to R_3' is chosen so that the average value of the feedback voltage over a half-cycle just equals $2\overline{\Delta E_t}$. It may be shown that then the two resetting cores II and III are reset the same amount during this second half-cycle. Then, during the third half-cycle they are both saturated during the same interval of time and the output voltages across R_3' and R_4' are equal but of opposite polarity. Only during the second half-cycle does a net voltage exist across R_3'. Also, during the third half-cycle $\overline{\Delta E_t}$ drops to zero in the operation of the complete amplifier shown in Fig. 2. Thus the output of both derivative amplifiers drops to zero during the fourth and succeeding 1/2-cycles until a change in E_t produces a derivative pulse. If the polarity of $\overline{\Delta E_t}$ is reversed, the operation is similar but the net output voltage exists across R_4' and hence its polarity is also reversed.

Coupling Derivative Amplifiers to Integrator

In coupling the output of the derivative amplifiers to the integrator input, two requirements must be kept in mind. First, the coupling circuit must be such as to insure freedom from interaction by the two derivative amplifiers. Hence the coupling circuit shown in Fig. 2 was decided upon. If the gating core of the positive derivative amplifier saturates, the voltage e_{ac}' is divided among the resistance of the load winding, the isolating resistor R_3 and the control winding of the resetting core of the integrator. A voltage is induced into the other control winding connected to the negative derivative amplifier. This induced voltage causes current to flow through R_4 and is then shunted through r_{10} so that it does not affect the operation of the negative derivative amplifier.

The second requirement to be kept in mind when designing the coupling circuit is maximum power transfer from the derivative amplifiers to the integrator. The coupling circuit can be reduced to the equivalent circuit in Fig. 5. The resistors R_3 and R_4 have the value R, the load-winding resistances of the derivative amplifiers have the value r and the input resistance of the integrator has the value R_x. For maximum power transfer from the voltage source e_{ac}' to the resistor R_x, the values of R and R_x are found to be

$$R=\frac{1}{\sqrt{2}}r \tag{27}$$

$$R_x=\frac{1}{2}r \tag{28}$$

The voltage drop across R_x is found to be

$$e_x=\frac{1}{2(2+\sqrt{2})}e_{ac}' \tag{29}$$

or about one-sixth of e_{ac}'. The power delivered to the integrator is found to be

$$P_x=\frac{E_x^2}{R_x} \tag{30}$$

Substituting from equations 28 and 29

$$P_x=\frac{(E_{ac}')^2}{4(3+2\sqrt{2})r} \tag{31}$$

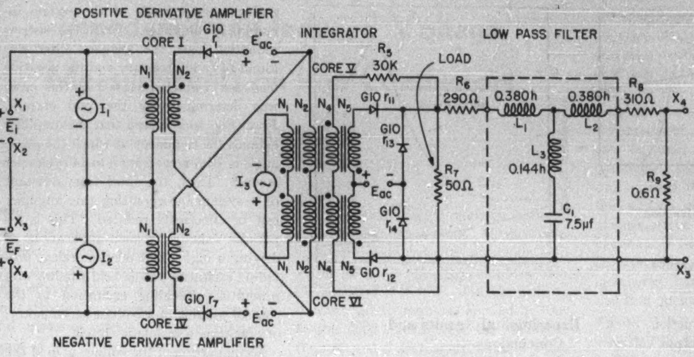
The maximum available power from the derivative amplifiers is

$$P_M=\frac{(E_{ac}')^2}{4r} \tag{32}$$

Hence the power delivered to the integrator is about one-sixth of the maximum available power.

It was pointed out in describing the operation of the integrator that the voltage across the load winding of the resetting core should not exceed e_{ac} if linearity between the output voltage and the integral of the input voltage is to be preserved. Making use of equation 29 for E_x, the relation between E_{ac}' and E_{ac} is thus determined to be

$$E_{ac}'=2(2+\sqrt{2})\frac{E_{ac}}{N} \tag{33}$$



CORE	TYPE	WINDING	TURNS
I	4168-S4	N_1	25
		N_2	3000
IV	4168-S4	N_1	25
		N_2	3000
V	4178-S4	N_1	50
		N_2	50
		N4	500
		N5	4000
VI	4178-S4	N_1	50
		N_2	50
		N4	500
		N_5	4000

Fig. 6. Complete amplifier incorporating half-wave derivative amplifiers

where N = turns ratio of the load winding to control winding of the integrator.

Low-Pass Filter

It was claimed in the section on circuit description that the inclusion of the low-pass filter in the negative feedback loop leads to greater sensitivity, stability, and speed of response. Greater sensitivity results since the low-pass filter attenuates second- and higher order harmonics of the power supply frequency included in the feedback voltage across R_9. Thus the magnitude of the alternating currents flowing through the signal source are greatly reduced.

Harmonics of the power supply frequency in the feedback voltage also increase the amount of variation of output voltage with temperature. Drift in the output of the amplifier shown in Fig. 2 is due primarily to variations in the leakage current through rectifiers, r_1, r_2, r_6 and r_7 with temperature. If the signal voltage and the feedback voltage were instantaneously equal, no voltage would exist backward across the rectifiers and hence the leakage current would be zero regardless of the resistance. Thus, although the backward resistances of the rectifiers are different functions of the temperature, the leakage currents would be zero and hence the output of the amplifier would be essentially independent of the temperature. However, harmonics in the feedback voltage make it impossible for the signal voltage to equal instantaneously the feedback voltage although their average values are equal. Thus part of each half-cycle voltage exists backwards across the rectifier in series with the load winding of the resetting core of the positive derivative amplifier. The same is true of the corresponding rectifier in the negative derivative amplifier. Thus, although the average values of these two voltages are equal when $\overline{E}_I = \overline{E}_F$, the backward resistance of the two rectifiers are not only unequal, but are generally different functions of the temperature. Hence, unequal leakage currents, the degree of inequality varying with the temperature, flow through the two rectifiers in the backward direction. This induces a temperature-dependent voltage in the control circuit which may either add or subtract from the feedback voltage causing the output voltage to be a function of the temperature. Thus by attenuating the harmonic components of the feedback voltage, the low-pass filter enhances the temperature stability of the amplifier.

The presence of the low-pass filter in the negative-feedback circuit gives rise to a time delay in the negative-feedback

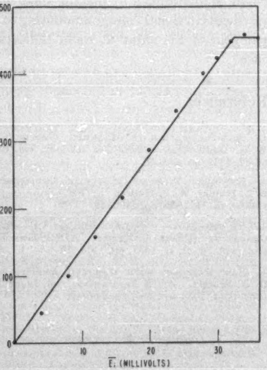


Fig. 7. Transfer characteristic

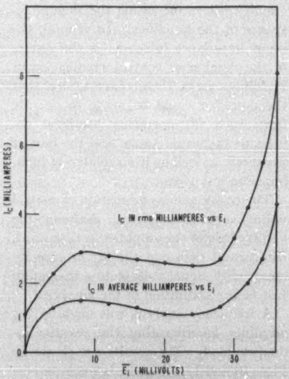


Fig. 8. Input characteristics

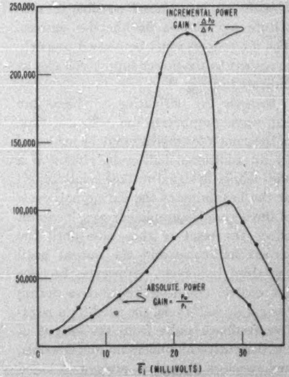


Fig. 9. Power gain

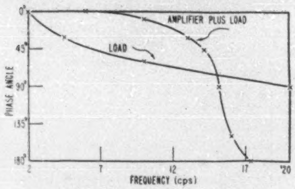


Fig. 10 (above). Phase function

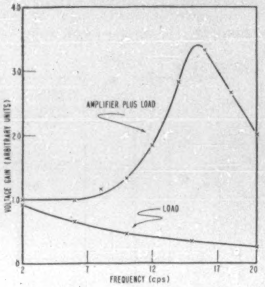


Fig. 11 (right). Gain function

voltage. After the introduction of a step of signal voltage, the output voltage overshoots its final quiescent value. Thus a derivative component is added to the output, increasing the over-all speed of response of any system in which the amplifier is incorporated.

A Simplified Amplifier Circuit

A simplified form of the 2-stage derivative-controlled amplifier is shown in Fig. 6. This simplification has been achieved by making use of half-wave derivative amplifiers eliminating two cores and six rectifiers. In addition, the necessity for isolating the output circuits of the two derivative amplifiers from one another has been eliminated. This is because interaction between the two circuits is not important since the half-wave amplifiers have a half-cycle dead time following the appearance of a derivative pulse in the output. Hence the two shunt rectifiers r_5 and r_{10} and the two isolating resistors R_3 and R_4 are not included in the circuit.

By making use of supermalloy cores rather than a rectangular hysteresis loop type it has been found possible to substitute a-c sources for the d-c sources. The a-c sources must be phased properly in respect to the power supply voltages as indicated in Fig. 6.

Because the derivative amplifiers are half-wave amplifiers, the response time of the complete amplifier may be increased by an additional half-cycle. Thus, if a step change in signal voltage is introduced at the beginning of the gating half-cycle of the derivative amplifier cores, it cannot affect the reset of the cores until the second half-cycle and the output until the third half-cycle. However, because the control circuit is inoperative during the gating half-cycle, the need for a negative feedback pulse from the output of the derivative amplifiers no longer exists, and resistors R_1 and R_2 are not included in the simplified circuit.

Experimental Results and Conclusions

The transfer characteristic of the simplified circuit is shown in Fig. 7. The load current is a linear function of the control voltage to a good approximation. Deviations of the transfer characteristic from a straight line were observed to be even smaller with the resistive-inductive load provided by the field winding of an 8-kva alternator.

The input characteristic of the simplified circuit is shown in Fig. 8. Both the average of the control current and the rms value of the control current are plotted as functions of the control voltage. The amplifier is stable in spite of the negative slope of the input characteristic in the range of signal voltage from 8 to 25 millivolts. This is because the rate of increase in negative feedback voltage is greater than the rate of decrease in current-resistance drop in the control circuit. This condition is satisfied in achieving 1-cycle response. For the circuit shown in Fig. 6, it was found that up to 5 ohms resistance could be inserted in the control circuit without increasing the time of response of the amplifier. In general, the input impedance increases as the square of the number of control-winding turns. It follows from these observations that the amplifier could be made to operate from such fast-operating low-level devices as the strain gauge and the barrier photocell, as well as from relatively high-impedance devices.

The power gain as a function of control voltage is shown in Fig. 9. Although the voltage gain of the amplifier is a constant, the current gain and the power gain increase with signal voltage to a maximum just before saturation of the amplifier.

A frequency analysis was made of the amplifier incorporating the resistive-inductive load. The results are shown in Figs. 10 and 11. Fig. 10 is a comparison of the phase angle of the load by itself and the phase angle of the system including the amplifier. With the resistive-inductive load, the low-pass filter was found to be unnecessary and the negative feedback voltage (delayed in this case) was determined by the load current. From Fig. 10 it is seen that the amplifier extends the frequency at which the phase angle is 45 degrees from 6 to 14 cycles per second. Thus, the total time constant of a system incorporating this amplifier can be effectively reduced. This is especially valuable in such applications as driving a high-speed teletype relay, providing excitation to the field winding of a generator, providing excitation to the control winding of a 2-phase servo motor, or operating a magnetic clutch.

A comparison of the voltage gain of the resistive-inductive load by itself and the system including the amplifier is shown in Fig. 11. It has been found that in practice the amplifier will operate with either supermalloy or Deltamax cores. The value of the maximum peremeability μ_m is a figure of merit for core materials in so far as their use in this amplifier is concerned.

It was pointed out in the description of the low-pass filter that the amplifier is essentially free from drift due to temperature variations. This is because the first stage operates as a derivative amplifier. Therefore, variations normally occurring in the parameters of this stage can only affect the time of response, and this to an insignificant amount. Some preliminary temperature tests on an amplifier similar to that shown in Fig. 6, but incorporating silicon junction rectifiers, were made. It was found that over the range from −25 to 70 degrees centigrade a constant load current of 100 milliamperes through a 70-ohm relay coil could be maintained by making small adjustments in signal voltage amounting to 1 per cent of the value at room temperature.

References

1. Self-Saturation in Magnetic Amplifiers, **W. J. Dornhoefer.** *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 835–50.

2. Response Time of Magnetic Amplifiers, **E. L. Harder, W. F. Horton.** *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1130–41.

3. On the Mechanics of Magnetic Amplifier Operation, **Robert A. Ramey.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1214–23.

4. Fast Response with Magnetic Amplifiers, **D. G. Scorgie.** *AIEE Transactions*, vol. 72, pt. I, 1953 (Jan. 1954 section), pp. 741–49.

◆

No Discussion

Impedance and Induced Voltage Measurements on Iron Conductors

W. F. MACKENZIE
ASSOCIATE MEMBER AIEE

THE purpose of this paper is to provide data and from these to draw conclusions on three questions involving the flow of 60-cycle alternating currents in iron or steel structures. First, the paper will attempt to provide a simple empirical means of calculating the a-c impedance of iron or steel members. The application of these data for protective grounding and fault calculations will be presented in a companion paper. Second, the paper will indicate the effect of joints and couplings in iron and steel structures on the flow of 60-cycle currents. The third question that this paper will attempt to answer is what effect the current flow on iron conduit surfaces will have on control conductors in the conduits. The application of these data will also be covered in a companion paper.

Impedance of Iron and Steel Members

The steel members on which impedance measurements were taken belong in two quite dissimilar groups. The first group consisted of I-beam type steel structural members varying quite widely in shape and in cross-sectional area from 5.88 to 206.25 square inches. The second group consisted of steel conduits varying in size from 1 to 4 inches.

The impedance of the members was measured by passing current through the steel and measuring the voltage drop along a known length. Test current was supplied by two 5-kva 240/10-volt transformers connected in parallel and measured with a 400-1 wound type of current transformer and portable ammeter. In the case of the structural columns, current connections were made 5 feet apart to the four edges of the columns by means of eight flexible *0000* leads fastened to the column with *C* clamps; see Fig. 1. Four voltage drops, one on each side of the column, were measured on the 2-foot section located midway between the current connections. These voltages were measured with a vacuum-tube voltmeter and a stabilized amplifier having a gain of 28.5 to 1. For the tests on conduit, the current was applied by four flexible *0000* leads held to the pipe by *C* clamps, and 5 feet apart. One voltage measurement was made on the 2-foot section located midway between the current leads, as indicated in Fig. 2.

Early in the tests, trouble was experienced by stray flux linking the voltmeter leads. The current in the iron products a magnetic field that may be divided into two parts. First, the field that is confined to the iron for the length of the applied current; and second, the field produced outside of the iron. For low saturations, almost all of the flux produced by the current will be confined to the iron surface and the error in the voltage caused by stray flux will be small. However, if a copper or a saturated steel member is tested, an appreciable amount of the flux produced by the current will exist outside the surface of the conductor. An undetermined amount of this flux will be cut by the voltage leads and will produce an induced voltage. This difficulty led to the use of a piece of 3/4-inch angle iron as a magnetic shield for making the two test-voltage measurements. The angle iron, as shown in Fig. 2, is insulated from the column or pipe by a thin layer of tape and encloses the voltage leads from the points of contact back to the center of the shield where the leads leave the shield as a twisted pair and connect to the vacuum-tube voltmeter terminals several feet away. The adequacy of this shield in shunting the magnetic field around the voltage leads was shown by a test made on a 1/4 by 2-inch copper bar on which the resistance, measured by applying 2,000 amperes and measuring the voltage drop on 2 feet, checked the calculated resistance within 3 per cent. The results of impedance measurements on columns and conduits appear in Table I.

To make the data useful for general calculations, Fig. 3 presents curves of the ratio of 60-cycle impedance to d-c resistance versus current density in amperes per square inch. An example of the application of these curves to calculate the electrical performance of an iron member is shown in the Appendix. The curves illustrate the effect of the shape of the member on the results and also indicate the effects of saturation. These curves can be used two ways. First, with reasonable accuracy the actual curves can be used to compute impedances of the various conduit sizes; second, an average curve can be used to estimate impedances on any iron configuration. In using data this way, it should be remembered that the effect of the shape of the iron member on the impedance is greater at low current densities, and the calculations will, therefore, be less accurate.

Effect of Joints on the Impedance

In the construction of a steel building framework the steel members are joined

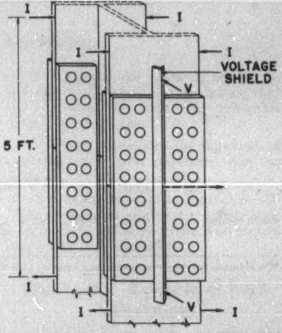


Fig. 1. Impedance measurement on column joint

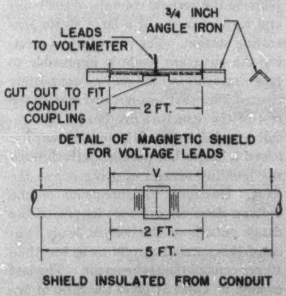

3/4 INCH ANGLE IRON
LEADS TO VOLTMETER
CUT OUT TO FIT CONDUIT COUPLING
2 FT.
DETAIL OF MAGNETIC SHIELD FOR VOLTAGE LEADS
V
2 FT.
5 FT.
SHIELD INSULATED FROM CONDUIT

Fig. 2. Impedance measurement on conduit coupling

Paper 54-469, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Middle Eastern District Meeting, Reading, Pa., October 5–7, 1954. Manuscript submitted July 6, 1954; made available for printing August 2, 1954.

W. F. Mackenzie is with the Pennsylvania Power & Light Company, Hazleton, Pa.

Table I. Results of 60-Cycle A-C Impedance Measurements on Columns and Conduits

Specimen Tested	Area, Square Inches	D-C Resistance, Microhms per Foot	Current, Amperes	Current Density	A-C Impedance, Microhms per Foot	A-C Impedance / D-C Resistance
8-inch wide flange I beam	5.88	13.8	810	138	*92	6.66
			1,540	262	76	5.50
			3,080	523	51	3.69
18-inch wide flange I beam	14.71	5.55	1,090	74.2	74	13.3
			2,112	144	60	10.8
Structural column *G62*	87.04	1.87	1,405	16.1	33.4	17.9
			2,580	29.6	29.6	15.8
Structural column *S56*	125.25	0.65	1,258	10.0	23.4	36.0
			2,360	18.8	24.7	38.0
Structural column *I47*	206.25	0.395	1,360	6.59	15.1	38.2
			2,560	12.4	16.2	41.0
1-inch conduit	0.53	154	200	378	625	4.07
			500	944	352	2.29
			1,000	1,890	213	1.39
1¹/₄-inch conduit	0.68	120	200	294	528	4.42
			500	735	329	2.76
			1,000	1,470	204	1.71
1¹/₂-inch conduit	0.79	103	200	253	511	4.96
			500	634	317	3.08
			1,000	1,270	197	1.92
2-inch conduit	1.03	79	200	194	435	5.52
			500	486	285	3.62
			1,000	970	186	2.36
3-inch conduit	2.31	35.2	40	17.3	285	8.1
			100	43.3	326	9.26
			200	86.6	330	9.38
			200	86.6	358	10.2
			400	173	262	7.44
			500	216	234	6.65
			600	260	212	6.02
			800	346	182	5.17
			1,000	433	157	4.47
			1,000	433	160	4.55
			1,200	520	142	4.03
			1,400	606	129	3.66
			1,600	693	117	3.32
			1,800	780	108	3.06
			2,000	866	99	2.81
			2,200	953	93.5	2.66
			2,400	1,040	87.5	2.49
			2,600	1,120	81.5	2.32
			2,800	1,210	76.5	2.18
			3,000	1,300	73.0	2.07
			3,200	1,380	69.0	1.96
			3,400	1,470	66.0	1.88
			3,600	1,560	63.5	1.81
			3,800	1,640	57.0	1.62
			4,000	1,730	55.0	1.56
			4,960	2,150	45.3	1.29
4-inch conduit	3.40	23.9	200	58.9	256	10.7
			500	147	193	8.10
			1,000	294	141	5.93

*On columns average voltage on four sides divided by current and proper constants.

Table II. Effect of Riveted Joints on the Impedance of Structural Columns

Column Number	Area,* Square Inches	Current Density	Ratio of Impedances of Joint to Column
S51	85.4	13.8	1.07
		25.6	1.01
S54	85.4	13.2	1.00
		25.0	1.03
G50	108.1	13.4	2.30
		24.0	2.56
G53	108.1	13.6	2.28
		24.4	2.34
G57	108.1	12.0	2.33
		21.8	2.32
G60	108.1	13.3	2.25
		24.3	2.20
S48	166	8.10	1.30
		15.2	1.25
S56	166	7.21	1.45
		13.7	1.42
I52	208.9	6.23	1.70
		11.7	1.54
I55	208.9	6.49	1.57
		11.7	1.41
I58	208.8	5.94	1.68
		10.6	1.60

*Cross-sectional area of column at joint.

by steel plates spanning the joint and are either riveted or welded to the member on both sides of the joint. Table II shows the results of tests made on riveted joints in the vertical columns of the boiler and turbine rooms of a large steam generating station. The procedure of erecting the structure made it impossible to take measurements on a completely riveted joint that was not tied in with the rest of the structure overhead. The result was that some of the test current followed a longer but parallel path through the building instead of going across the joint. However, measurements of the voltage drop on a 2-foot length of the flange outside of the current leads indicated this parallel current to be less than 10 per cent of the test current. The last column in Table II gives the ratio of impedances of a riveted joint to a solid column of the same shape and cross-sectional area. The results are quite consistent and indicate that the impedance of the joints are not significantly higher than the continuous columns.

Table III gives the results of tests made on conduit couplings with various sizes of conduit and three different types of joint sealer. The impedance figures in this table indicate that:

1. If the joints are properly threaded or have the threads cut deeply enough to run the coupling on by hand, the final joint will be satisfactory, provided the ends of the conduit are firmly butted together.

2. As long as the conduit ends are touching, very little current passes through the coupling. Therefore, the impedance of the joint is practically the same as the straight conduit and is unaffected by the type of sealer used.

Conduit Voltage Changed by Path of Current

To learn more about the performance of iron circuits carrying current, the path of the current relative to the conduit was changed. As indicated in Fig. 4, a 2-foot voltage drop on the outside surface of the pipe was measured with three different types of current flow. The first condition, similar to previous conduit

Table III. Impedance of Conduit Couplings with Various Sealers Compared to Straight Conduit at 1,000 Amperes, 60 Cycles

Conduit Size	Straight Conduit, Microhms per Foot at 1,000 Amperes	Type of Joint Sealer, Microhms per Foot at 1,000 Amperes: No Sealer	Red Lead	Zinc Chromate	No Oxide Grease
1 inch	214	214	206	214	216
1¹/₄ inches	202	202	200	207	208
1¹/₂ inches	199	192	194	197	205
1¹/₂ inches with loose threads	176	185	178	184	185
2 inches	185	188	185	188	188
2 inches with loose threads	185	192	185	188	187
3 inches	160	232*	204*	200*	
				192	178
4 inches	138	142	142	142	142

*Coupling was too small to permit the ends of the conduits to meet.

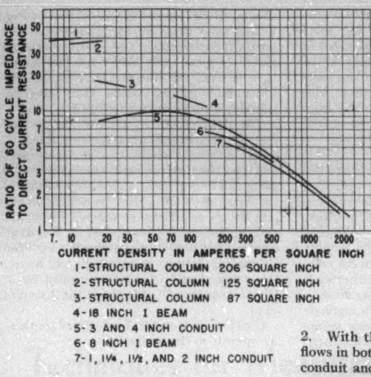


Fig. 3 (left). Iron conductor impedance ratio as a function of current density; for detail of voltage shield see Fig. 2

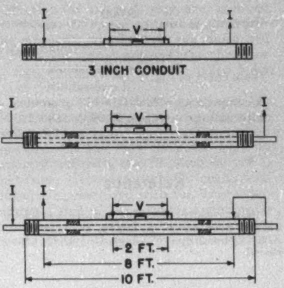


Fig. 4 (right). Voltage measurements on conduit surface for three different types of current flow

tests, carried the current out on the conduit and provided a remote return. The second setup carried the current out on a 3/4-inch copper tube inside, but insulated from the conduit with a remote return. The third setup carried the current out, as in the second case, but returned on the conduit. The results of these tests appear in Table IV. These results indicate that the current on the outer surface of the conduit is the same in the first and second cases and is zero in the third case. These results confirm the results obtained in reference 1. Reference 1 also provides a good explanation of these results as follows:

1. With the first condition, the current flows only in the outer surface of the conduit and the impedance of the circuit is approximately equal to the conduit impedance.

2. With the second condition, the current flows in both inner and outer surfaces of the conduit and the impedance of the circuit is approximately equal to twice the conduit impedance.

3. With the third condition, the current in the outer surface of the conduit induced by the copper tube inside the pipe is cancelled by current returning on the conduit and the resultant conduit current, therefore, is on the inner surface of the conduit. The impedance of this circuit is approximately equal to the conduit impedance.

Table IV. Voltage Drop Along the Outer Surface of a 3-Inch Conduit for Different Types of Current Circuits

	Millivolt Drop on 2-Foot Surface		
Current, Amperes	Current on Conduit, Return Remote	Current inside Conduit, Return Remote	Current inside Conduit, Return on Conduit
100	65.2	71.8	
200	132	126	0
400	210		0
600	255	247	0
800	291		0
1,000	314		0
1,200	342	330	0+
1,400	362		0.3
1,600	374		0.3
1,800	391	388	0.6
2,000	399		0.6
2,200	411		0.9
2,320		427	
2,400	419		1.1
2,600	425		1.1
2,800	430		1.4
3,000	439		2.0
3,200	442		2.0
3,400	448		2.8
3,600	456		3.4

Induced Voltages on Conductors in Conduits Carrying Fault Current

If conduits are used to carry ground currents during faults, the effect of these ground currents on control conductors should be determined. This condition was set up experimentally on 30-feet of a 3-inch conduit, as shown in Fig. 5. The various voltages measured are shown in Table V. The results of this test indicate that while appreciable voltages appear along the conduit a similar voltage is induced on the control conductor in the conduit. Therefore, if the control cable is connected to the conduit at one end, the voltage between conduit and conductor at the other end will only be a small percentage of the total drop along the conduit. The neutralizing effect of the conduit will be decreased as the saturation of the iron is increased. In this respect, the data are limited to a current of 1,400 amperes on 3-inch conduit or a density of 606 amperes per square inch.

Table V. Induced Voltages on a Conductor in a Conduit Carrying Current*

Current on Conduit, Amperes	V_1, Volts	V_2, Volts	V_3, Volts	V Pipe,† Volts	$(V_3/V_p)\times100$, Per Cent
100	1.16	1.16	0+	0.95	
200	2.20	2.20	0+	1.92	
400	3.55	3.50	0.02	3.04	0.66
600	4.60	4.52	0.03	3.70	0.81
800	5.70	5.62	0.05	4.22	1.19
1,000	6.60	6.55	0.07	4.55	1.54
1,200	7.46	7.35	0.10	4.96	2.02
1,400	8.40	8.30	0.13	5.25	2.48

*See Fig. 5 for diagram of connections.

†This voltage is obtained by multiplying the impedance of a 3-inch conduit in column 6 of Table I by the length of 29 feet and the proper current. Since the data in Table I were obtained using shielded voltage leads, this voltage is not affected by stray flux.

Conclusions

1. By means of the curves of Fig. 3, the impedances of various iron conductors may be determined for fault and protective grounding calculations.

2. The data indicate that the impedance of joints in these iron structures may be neglected.

3. If conduits are used as grounding conductors, potential gradients resulting from fault currents will be neutralized on control conductors in these conduits.

Appendix. Example of Fault Calculation

An example of a fault calculation using the curves in Fig. 3. What voltage will appear along 10 feet of a 36-square-inch column if it is carrying the total available short-circuit current of a 480-volt load center of 18,000 amperes?

The resistivity of iron is 6.79 microhms per 1-inch cube.

$$\text{D-c resistance} = \frac{6.79\times12\times10}{36} = 22.6 \text{ microhms}$$

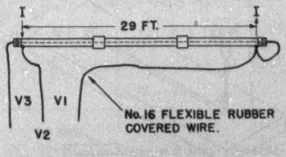


Fig. 5. Induced voltage on control conductor on conduit carrying current

$$\text{Current density} = \frac{18{,}000}{36} = 500 \text{ amperes per square inch}$$

$$\text{From curve of Fig. 3 } \frac{\text{a-c impedance}}{\text{d-c resistance}} = 4$$

$$\text{A-c impedance} = 22.6 \times 5 = 113 \text{ microhms}$$

$$\text{Alternating voltage} = 18{,}000 \times 113 \times 10^{-6} = 2.4 \text{ volts}$$

Reference

1. Iron Conduit Impedance Effects in Ground Circuit Systems, A. J. Bisson, E. A. Rochau. *AIEE Transactions*, vol. 73, pt. II, July 1954, pp. 104–07.

◆

Discussion

William Deans (I-T-E Circuit Breaker Company, Philadelphia, Pa.): I would like to know whether impedance or something approaching a-c resistance of the conductor is found by the method described. By the apparatus shown in Fig. 2, the effect upon the measured voltage drop of magnetic flux caused by current in the conductor is substantially eliminated. It is this magnetic flux that gives rise to the reactive component of the impedance drop. This leaves only the a-c resistance component.

Details of the argument are as follows: In this argument, it is assumed that the current in the voltmeter circuit is so small as to be negligible.

In Fig. 6 A and B are two points on the same stream line of current flow in conductor C. Points A and B are connected to the voltmeter V_M' by wire $A'B'$ parallel and close to current filament AB and by a twisted pair of wires. The voltmeter therefore measures the voltage loop in $BAA'B'B$.

Consider first the extreme case, where wire $A'B'$ and current filament AB are coincident. The magnetic field caused by the current in conductor C affects wire $A'B'$ and current filament AB in exactly the same way. That is, voltages AB and $A'B'$ due to this flux are identical; they are equal and in phase in the stated directions. The voltmeter sees the voltage in the loop; it therefore sees voltage AB minus voltage $A'B'$. Now, voltage AB is the phasor sum of the a-c resistance drop and the reactance drop. And the reactive drop is produced by the magnetic flux in conductor C and is equal to voltage $A'B'$. The voltmeter reading is therefore the impedance drop minus (phasorially) the reactive drop in the conductor; it is the a-c resistance drop.

As wire AB is displaced from filament AB, some of the magnetic flux caused by the current will pass through the loop and some part of the reactive drop will be included in the voltmeter reading. To include all of the reactance drop, i.e., to have the voltmeter read the whole impedance drop, wire $A'B'$ must be "at infinite distance" from filament AB and connections AA' and BB' must be very long and perpendicular to filament AB.

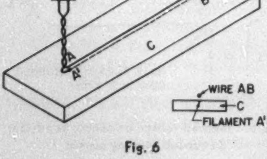


Fig. 6

The angle iron of Fig. 2 shunts flux to the outside which otherwise would pass through the loop. In effect, it brings the measuring wire closer to the conductor. It, therefore, serves to reduce the effect on the voltmeter of the reactive drop. Incidentally, it matters not whether the twisted pair to the voltmeter is taken from the center, either end, or any other place in the measuring wire.

It appears that by the author's method, the total impedance has not been measured but, instead, something which closely approximates the a-c resistance.

The curves of Fig. 3 of the paper are consistent with this analysis. Stated briefly, the argument is as follows.

The conductors are of magnetic material and, therefore, subject to saturation. Because of nonuniform current density over the cross section, saturation varies also over the cross section. As saturation increases, the skin effect behavior of the conductor approaches that of a conductor of nonmagnetic material. The skin effect ratio improves, i.e., R_{a-c}/R_{d-c} decreases. So it is to be expected that, as current density and, therefore, saturation increase, the a-c resistance decreases.

L. E. Rinker (Leupold and Stevens Instruments, Inc., Portland, Oreg.): Study of the curves shown in Fig. 3 appears to establish a relationship between the area of the conductor and the character of the curve. Curve 3, however, appears to depart from the trend shown on all the other curves. Six of the curves would appear to have inflection points flowing between 30 and 100 amperes per square inch. The inflection point for curve 3 obviously is very much lower. I wonder if one curve cannot be constructed which would be sufficiently accurate for all practical purposes as applying to all iron conductors, regardless of size or shape. Curve 5 probably is representative for this universal application. I suspect that curves 1, 2, and 3 are in error because of the method of obtaining the d-c resistance when conducting the tests.

It would be entirely possible to measure the resistance directly opposite each other in the structural member. It is quite apparent also that no relationship could be established between this d-c resistance and the a-c impedance offered to a current flowing longitudinally in the structural member. As the points of resistance measurement are displaced longitudinally a current-impedance drop will occur and it will be possible to establish a relationship between the d-c resistance and the a-c impedance for that particular structural member. For a large cross-section area conductors, the measurement points should be a long way apart to obtain accuracy. A 3-inch conduit has a cross-section area of approximately 1 square inch and curve 5 was constructed from data where the points were 2 feet apart. If the points of measurement for the structural column of 206 square inches, shown in Fig. 3, had been made in like proportion, the points of measurement should have been displaced longitudinally about 400 feet apart. It is believed that the test data for the 206-square-inch column would then have resulted in a curve similar to Fig. 5.

The data presented are for one specific case, i.e., resistance measurements taken at points displaced longitudinally 2 feet apart and, therefore, all problems should be set up using this limitation. The example given in the Appendix should then be revised to obtain the d-c resistance at points 2 feet apart and by proper usage of the curves a voltage calculation should be obtained. For the alternating voltage at points 10 feet apart, this value should then be multiplied by 5. It undoubtedly will differ from the 2.4 volts cited.

Conclusions 2 and 3 should be of particular interest to the profession.

W. F. Mackenzie: I wish to thank the discussers for their interest in the paper. The question raised by Mr. Deans is a good one as it brings out a point not indicated in the paper. The substance of his argument is entirely correct, although I feel that his conclusion needs some clarification.

For the purpose of this analysis, the total impedance of a magnetic conductor will be divided into three parts:

1. The a-c resistance.
2. The reactance resulting from flux present inside the surface of the magnetic conductor.
3. The reactance resulting from flux present outside the surface of the magnetic conductor.

As Mr. Deans points out, the third part of the total impedance is eliminated by using the angle-iron shield. The total impedance may be measured by bringing the voltage lead away from the specimen perpendicular to its axis.

Stated simply, the impedance measured in the paper and indicated in the curves of Fig. 3 is the phasor sum of the a-c resistance and the internal reactance.

Experimental values of total impedance, although just as easy to obtain, do not lend themselves to a general application because the external impedance depends entirely on the geometry of the circuit. The external impedance for simple configurations may

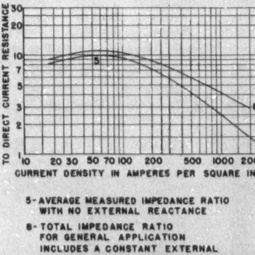


Fig. 7. Iron conductor impedance ratio as a function of current density

be calculated from classical equations.

If the application requires the addition of the external impedance, the question of the phase angle at which to add it arises. Referring to the data on 3-inch conduit contained in reference 1, the resistance and reactance components of total impedance are presented for current values of 250, 500, and 1,000 amperes. From the dimensions given for the test circuit in this reference, the external reactance may be calculated to be 72 microhms per foot. A subtraction of this reactance from the total reactance will give the internal reactance. The phasor sum of the resistance and internal reactance components yields an impedance whose phase angle is practically constant at 30 degrees for the three current values of 250, 500, and 1,000 amperes. In Fig. 7, an external reactance of 72 microhms per foot is added to curve 5 of Fig. 3 to produce curve 8.

For a particular application, if the external reactance differs markedly from the arbitrary value of 72 microhms per foot, the total impedance may be obtained either by interpolation between curves 5 and 8 or by vector addition of the external reactance to the impedance of curve 5 taken at a phase angle of 30 degrees.

As indicated by Mr. Rinker, curve 5 seems to be representative for a general application and accordingly curve 8, including a correction for external reactance, is based on it. By the same method employed in the Appendix, the external reactance will increase the impedance and, therefore, for general application, the ratio taken from curve 8 will be 6.

In answer to Mr. Rinker's question on the d-c resistance measurement for the large columns, these figures were not obtained by measurement but were computed from an assumed resistivity of 6.79 microhms per 1-inch cube. However, the small longitudinal distance as compared with the large cross-sectional area of the columns undoubtedly does contribute in some manner to the apparent errors in curves 1, 2, and 3 of Fig. 3.

Techniques for Measuring Cascaded Self-Saturating Magnetic-Amplifier Performance

HENRY KAPLAN
NONMEMBER AIEE

GERALD WOLFF
ASSOCIATE MEMBER AIEE

THE rapid development of magnetic amplifiers in recent years has led to the spending of much effort on their analysis and design criteria. Despite the progress made, a major part of the design must still be carried out experimentally.

Determination of amplifier performance by the ammeter, voltmeter, and wattmeter methods becomes difficult since almost all current and voltage wave shapes encountered are nonsinusoidal. Moreover, the effects of changes in various circuit parameters are not easy to analyze. These difficulties are especially apparent in the half-wave bridge-type circuits where three or four stages may be cascaded, and the signals after the first stage are nonsinusoidal. This paper presents experimental techniques which utilize the basic concept of core flux density for determining magnetic-amplifier performance.

Core Flux Density as a Measure of Magnetic-Amplifier Performance

Flux density in the core may be considered as the fundamental quantity in a magnetic amplifier, since it directly determines the magnitude of the output. A brief derivation of this relationship will be found at the end of this section. Several authors[1-3] have discussed flux density measurements in the course of determining dynamic flux loops and control characteristics. However, rather elaborate equipment and precise calibration techniques are required, and the results cannot be readily applied as performance criteria.

Fig. 1 shows dynamic hysteresis loops for various control signals. B_f is the "firing" flux density,[4] and for rectangular core materials may be assumed to coincide with the saturation flux density. For less rectangular core materials, such as permalloy, B_f will be slightly lower. B_0 is the value of flux density at the time t_0, i.e., when the supply voltage passes through zero just preceding the conducting half-cycle. B_0 varies from the minimum value of the loop for d-c control to the intersection of the loop with the B-axis (Fig. 1) for control signals applied only during the nonconducting cycle. The net change of flux density in a core caused by a given control signal is expressed by $B_f - B_0$. This quantity, or the more useful nondimensional ratio B_0/B_f, may be obtained experimentally. One method is that of reference 4, which assumes rectangular core material, perfect rectifiers, sinusoidal applied voltage, and load resistance small compared to unsaturated coil reactance but large compared to saturated coil reactance.

Using this method, an expression for B_0/B_f may be derived as follows

$$E_m \sin \omega t = N\frac{d\Phi}{dt}10^{-8} \quad (1)$$

$$\Phi_f - \Phi_0 = (E_m 10^8/\omega N)(1-\cos\theta_f) \quad (2)$$

$$B_f - B_0 = (E_m 10^8/\omega NA)(1-\cos\theta_f) \quad (3)$$

$$B_0/B_f = 1-(E_m 10^8/\omega N B_f A)(1-\cos\theta_f) \quad (4)$$

$$B_0/B_f = 1-K(1-\cos\theta_f) \quad (5)$$

where

Φ_f = core flux at saturation (firing)
Φ_0 = core flux at time t_0
E_m = maximum value of supply voltage
θ_f = firing angle
A = cross-sectional area of core
N = number of turns on anode winding
$K = E_m 10^8/\omega N B_f A$.

$B_f A$ or ϕ_f may be determined by measuring the rms voltage E just sufficient to saturate the core wound with N

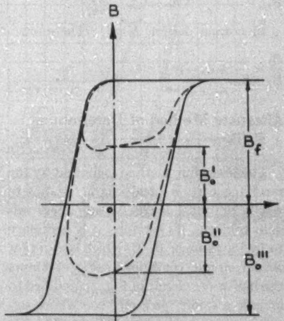


Fig. 1. Typical major and minor dynamic-hysteresis loops for rectangular-type core materials

---

Paper 54-325, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11-15, 1954. Manuscript submitted June 10, 1953; made available for printing June 23, 1954.

HENRY KAPLAN and GERALD WOLFF are with the Ford Instrument Company, Division of the Sperry Corporation, Long Island City, N. Y.

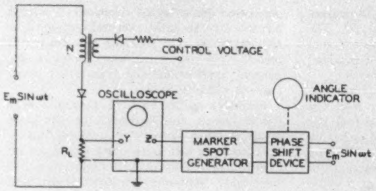


Fig. 2 (left). Circuit for obtaining B_0/B_f by firing angle method

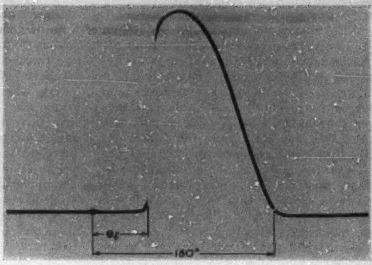


Fig. 3 (right). Oscillogram of output wave form as used in determination of firing angle

turns. Thus $B_fA = \dfrac{E\times 10^8}{(4.444f)N}$, where f is the frequency of the applied voltage. The firing angle (θ_f) is measured with the aid of an oscilloscope and a marker-spot generator, (Fig. 2), the phase of which can be continuously shifted and accurately calibrated. By viewing the voltage across the load and injecting the marker spot on the z-axis of the oscilloscope, the angle at which firing takes place can be easily measured by shifting the marker spot from the predetermined zero to the firing point, which is well defined for rectangular core materials. Fig. 3 shows a typical oscilloscope pattern. The marker spot is at the zero point.

It may be stated that the output of a magnetic amplifier utilizing square-loop core material is directly proportional to the area under the load voltage curve from the firing point to extinction. B_0/B_f can be related to this area, and therefore to the output, as follows

$$\text{Area} = \frac{1}{\omega}\int_{\theta_f}^{\pi} E_m \sin\theta\, d\theta$$

$$= \frac{E_m}{\omega}(1+\cos\theta_f)$$

Combining with equation 5

$$\frac{B_0}{B_f} = 1 - K + K\left(\frac{\omega}{E_m}\text{ Area} - 1\right)$$

In normal design, $K=1$. Therefore

$$\frac{B_0}{B_f} = \frac{\omega}{E_m}\text{ Area} - 1$$

Alternate Method of Determining B_0/B_f

The preceding method is limited by the need for certain ideal characteristics to which practical magnetic amplifiers seldom conform. Therefore, it is necessary to find a value of B_0/B_f which is valid for actual circuit conditions. Fig. 4 shows the half-wave circuit of Fig. 2, modified to include a small "pick-off" coil whose signal is integrated and amplified. The following symbols are used to derive an expression for B_0/B_f:

$E_m \sin\omega t$ = supply voltage
N = number of turns in anode winding
N_i = number of turns in integrator pick-off coil
$\Phi(t)$ = flux in core
Φ_f = core flux at saturation
Φ_0 = core flux at time t_0
Φ_A = average value of core flux
B_f = core flux density at saturation
B_0 = core flux density at time t_0
$e_i(t)$ = instantaneous input voltage to integrator amplifier
$e_0(t)$ = instantaneous output voltage of integrator amplifier
K' = constant of integrator amplifier
T = period for one cycle of supply voltage
A = cross-sectional area of core

B_0/B_f can now be derived as follows

$$e_i(t) = N_i\frac{d\Phi(t)}{dt}10^{-8} \qquad (6)$$

$$\int_{t_1}^{t_1+T} e_i(t)dt = N_i\int_{t_1}^{t_1+T} 10^{-8}\frac{d\Phi(t)dt}{dt}$$
$$= N_i[\Phi(t_1+T)-\Phi(t_1)]10^{-8} \qquad (7)$$

Dividing by T

$$\frac{1}{T}\int_{t_1}^{t_1+T} e_i(t)dt =$$
$$(N_i/T)[\Phi(t_1+T)-\Phi(t_1)]10^{-8} \qquad (8)$$

but in the steady state

$$\Phi(t_1+T) = \Phi(t_1)$$

and

$$\frac{1}{T}\int_{t_1}^{t_1+T} e_i(t)dt = 0 \qquad (9)$$

The average value of $e_i(t)$ in the steady state is thus equal to zero. By a Fourier series expansion it is possible to show that

$$\frac{1}{T}\int_{t_1}^{t_1+T} dt'\int e_i(t)dt = 0$$

Therefore, since $e_0(t)\underline{\Delta}K'e_i(t)dt$, with the constant of integration eliminated by the a-c characteristic of the integrator, the average value of $e_0(t)$ in the steady state is also equal to zero. Then

$$e_0(t) = K'\int e_i(t)dt = K'N_i10^{-8}\int\frac{d\Phi(t)}{dt}dt \qquad (10)$$

$$= K'N_i10^{-8}[\Phi(t)+C] \qquad (11)$$

and

$$\frac{1}{T}\int_{t_1}^{(t_1+T)} e_0(t)dt = (K'N_i10^{-8}/T)\times$$
$$\int_{t_1}^{(t_1+T)}\Phi(t)dt + K'N_i10^{-8}C \qquad (12)$$

therefore

$$0 = \frac{1}{T}\int_{t_1}^{(t_1+T)}\Phi(t)dt + C \qquad (13)$$

$C = -\Phi_A$ or average flux in core

$$e_0(t) = K'N_i10^{-8}[\Phi(t)-\Phi_A] \qquad (14)$$

$$\Phi(t) = e_0(t)/K'N_i10^{-8} + \Phi_A \qquad (15)$$

at the time t_0

$$\Phi(t) = \Phi_0 = [e_0(t_0)/K'N_i10^{-8}] + \Phi_A$$

Also

$$\Phi_f = [e_{om}(t)/K'N_i10^{-8}] + \Phi_A$$

where $e_{om}(t)$ is the maximum instan-

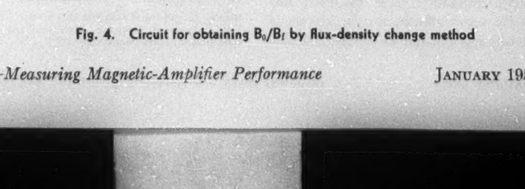


Fig. 4. Circuit for obtaining B_0/B_f by flux-density change method

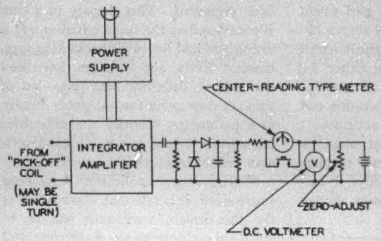


Fig. 5 (left). Circuit of a magnetic - amplifier performance analyzer

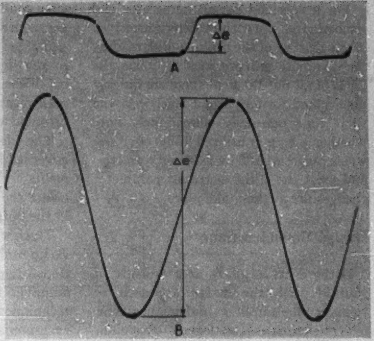


Fig. 6 (right). Oscillograms of flux density variation in core of a half-wave self-saturating circuit

A—With no control

B—With full control

taneous integrator output voltage for a given control signal. Thus

$$\Phi_0-\Phi_f=\frac{e_0(t_0)-e_{om}(t)}{K'N_t10^{-8}} \tag{16}$$

$$B_0-B_f=\frac{e_0(t_0)-e_{om}(t)}{K'N_tA10^{-8}} \tag{17}$$

$$B_0/B_f=1-\frac{e_0(t_0)-e_{om}(t)}{K'N_tB_fA10^{-8}} \tag{18}$$

$$B_0/B_f=1-K''\Delta e \tag{19}$$

where

$$K''=10^8/K'N_tB_fA$$

$$\Delta e=e_0(t_0)-e_{om}(t)$$

B_fA may be determined as in the previous method, and K' and N_t are known constants. Therefore the only quantity to be obtained experimentally is Δe. Since this is the difference between two instantaneous values of the integrator output voltage at defined points of 1 cycle, the value of Δe is found by viewing the output wave form of the integrator on a calibrated oscilloscope and measuring the amplitude difference between $e_0(t_0)$ and $e_{om}(t)$. $e_{om}(t)$ is the maximum value of the wave form and $e_0(t_0)$ is indicated by the marker-spot generator when set at t_0 on the time base.

In most cases $e_0(t_0)$ is so close to the minimum value of the wave form that a peak-to-peak reading meter may be substituted for the oscilloscope and Δe read directly. A schematic for an instrument utilizing such a meter is shown in Fig. 5. This self-contained, direct-reading unit eliminates the need for a calibrated oscilloscope and is much more sensitive to low-level voltage difference readings. The output of the integrator amplifier is passed through a d-c restorer and then rectified. The zero-adjust is set so that there is no deflection of the center-reading meter pointer. The voltmeter reads the Δe for this signal level. As the signal is varied, the change in Δe is measured by the center-reading meter.

If the control signal is such as to cause current flow in the load circuit during some part of the reset cycle, $e_0(t_0)$ may differ enough from the minimum value to require a correction in the meter readings if high accuracy is desired. However, this is never a desirable mode of operation. For a continuous picture of what is taking place in a core, the oscilloscope remains the ideal device. Fig. 6(A) and (B) are typical traces of the flux-density levels at fully saturated and unsaturated conditions, respectively.

An experimental verification of this method is demonstrated by the graphs of Fig. 7. By the use of vacuum-tube rectifiers, sinusoidal input, and high ratio of unsaturated coil impedance to load resistance, the ideal conditions for the firing-angle method were approximated. B_0/B_f was obtained for various circuit parameters in accordance with the two methods outlined. The resulting plots (Fig. 6) show that there is a close correlation. As would be expected, the values of B_0/B_f as calculated by the firing-angle method are slightly greater, since the core actually "fires" at a flux-density level which is somewhat less than saturation.

Definition of Gain

While curves such as those of Fig. 6 are often useful, being the actual dynamic characteristics for d-c control, it is in systems where several amplifiers are cascaded that the utility of a quick and easy method of measuring B_0 becomes most useful. At the present writing, the gain of a magnetic amplifier stage in a cascaded system is still undefined. Gains have been given in terms of power or of average rms, or peak values of current or voltage. For various special conditions the use of any one of these quantities may give adequate results, but a more accurate and universal definition of gain is suggested by the fundamental quantity of the magnetic amplifier, the core flux density. In cascaded stages, the flux-density change in a core is caused by the flux-density change in the core of the previous stage. Formalizing this fact, the gain G between two stages may be expressed as

$$G=\frac{\left(\Delta\dfrac{B_0}{B_f}\right)_{\text{stage 2}}}{\left(\Delta\dfrac{B_0'}{B_f'}\right)_{\text{stage 1}}}=\frac{\Delta e_2-\Delta e_1}{\Delta e_2'-\Delta e_1'}\,\frac{B_f'A'}{B_fA}$$

If the cores in both stages under consideration are identical, the gain becomes

$$G=\frac{\Delta e_2-\Delta e}{\Delta e_2'-\Delta e_1'}$$

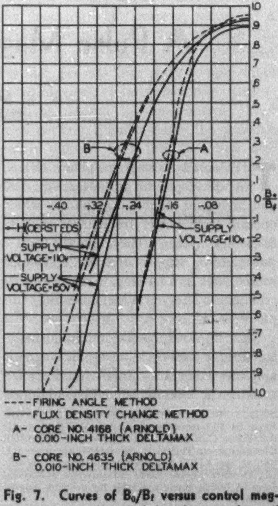


Fig. 7. Curves of B_0/B_f versus control magnetization as obtained by firing-angle and flux-density methods

For a specified signal, numerator and denominator can each be determined by a single reading of the peak-to-peak reading meter mentioned previously. Thus the gains of all but the final stage are quickly determined for any signal level. At the same time one may immediately note whether a core is utilized fully over the whole range, whether it is biased correctly, and what the effect of circuit parameter changes are as they take place.

Design Considerations

If the criterion of B_0 as a measure of gain is used, the design approach to vacuum-tube amplifiers and magnetic amplifiers becomes analogous. Vacuum-tube amplifiers are commonly divided into three distinct parts, input stage, voltage amplifier, and output stage, based on the fact that the vacuum tube is essentially a voltage-amplifying device. Similarly, a cascaded magnetic amplifier may be divided as follows: input stage B_0 amplifier, output stage.

In general, each part requires a different design approach. The design of the input stage is usually determined by time-constant considerations, and would consequently be a function of source impedance. The wave forms encountered in the first stage are usually either d-c or sinusoidal and therefore calculations with regard to input stage operation can be made with a fair degree of accuracy.

The design of the output stage is governed by the load and power requirements. From these it is possible to calculate the maximum flux density change in the output core.

Thus, with an input stage designed and B_0 known for the given output specifications, the gain requirement of the B_0 amplifier is determined. By using the techniques discussed in this paper, gain characteristics of the B_0 amplifier stages can be obtained by direct measurement, thereby allowing the designer a rapid check on the adequacy of his design, the amplification of each stage, and the effects of various circuit changes.

Summary

An experimental procedure for the analysis of the performance of cascaded self-saturated magnetic amplifiers has been presented. Flux density in a core is considered as the basic criterion and a simple method for easily measuring flux-density levels and changes has been derived. A definition was suggested of gain per stage, with the use of flux density as a parameter, whereby the difficulties inherent in the use of nonsinusoidal voltages and currents could be avoided.

While practical applications have been emphasized, it is felt that observation of the flux-density variations can be of much aid not only in the improving and developing operating circuits but also in obtaining a better understanding of the fundamentals of magnetic amplifier action.

References

1. Predetermination of Control Characteristics of Half-Wave Self-Saturated Magnetic Amplifiers, **Henry Lehmann.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 2097–2103.

2. Dynamic Hysteresis Loop Measuring Equipment, **H. W. Lord.** *AIEE Transactions*, vol. 71, pt. I, Sept. 1952, pp. 269–72.

3. Evaluation of Core Materials for Magnetic Amplifiers—I and II, **D. C. Dieterly.** *Electrical Manufacturing*, New York, N. Y., vol. 50, Jan. 1953, p. 68; Feb. 1953, p. 124.

4. Self-Saturation of Magnetic Amplifiers, **W. J. Dornhoefer.** *AIEE Transactions*, vol. 68, pt. II, 1949, pp. 835–50.

The Biased Rectifier Amplifier—A Pulse Magnetic Amplifier

R. E. MORGAN
MEMBER AIEE

THE biased rectifier amplifier (BRA) was developed to provide small, fast and high-gain magnetic amplifiers which operate from power frequencies such as 60 or 400 cycles. They operate from a pulse power supply and use a voltage source, such as a battery, in series with its rectifiers. They provide push-pull output with one core.

The BRA reduces the minimum reactor size for a 60-cycle push-pull magnetic amplifier stage to a small fraction of a cubic inch; see Fig. 1. The minimum reactor size of a conventional magnetic amplifier is limited in most applications by a minimum wire size of about 0.002 inch diameter and a minimum supply voltage of about six volts rms (8.5 volts peak). A peak supply voltage E_p of more than 8.5 volts is usually used so that it will be much larger than the rectifier forward voltage drop e_r of about 1 volt. The pulse power supply provides a peak voltage E_p of more than 10 volts and yet the supply voltage averaged over a half-cycle is less than 0.001 volt. Pulse lengths of less than 3 microseconds and over 10 milliseconds have been used. The most commonly used pulse length is from 20 to 200 microseconds. The pulse repetition is commonly equal to the sine-wave frequency, which may be any frequency satisfactory for magnetic amplifiers. The BRA can provide push-pull output with one core and one gate winding (power winding) or provide full-wave output with one core and one gate winding, whereas normally two cores are required. The pulse power supply[1] is magnetic, using saturable reactors, or transformers, about the size of the minimum size reactor for conventional magnetic amplifiers. One pulse supply can service many BRA stages or amplifiers.

The BRA is a magnetic amplifier for amplifying control-type signals from about 10^{-12} to 10^{-3} watt. The BRA has

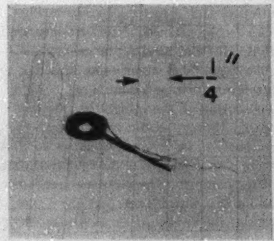


Fig. 1. BRA reactor for push-pull output at 60 or 400 cycles

Paper 54-364, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 14, 1954; made available for printing August 10, 1954.

R. E. Morgan is with the General Electric Company, Schenectady, N. Y.

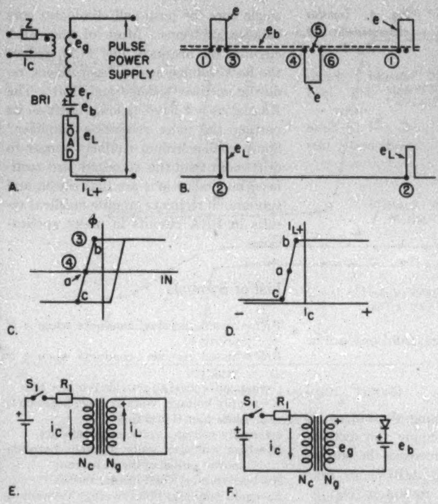


Fig. 2. Half-wave BRA

A—Circuit
B—Wave shape
C—Hysteresis loop
D—Transfer characteristics
E and F—Equivalent circuit between t_3 and t_4 with and without e_b

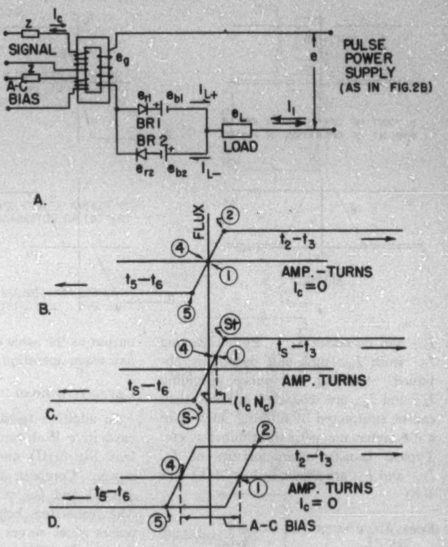


Fig. 3. Push-pull output

A—BRA push-pull circuit
B and C—Simplified magnetization curve without a-c bias with $I_c=0$ and I_c at $1/2\ \varphi_{s+}$ respectively
D—Hysteresis loop with a-c bias

similar gain and speed of response and uses the same core materials as conventional magnetic amplifiers.[2] The small reactor, Fig. 1, makes the BRA ideal for half-cycle response time per stage. The efficiency of its power supply, about 10 per cent, makes the BRA undesirable for outputs greater than 1 watt. This paper describes only a few of the many possible circuits for the BRA. Flux and ampere turns are normally used in this paper to simplify the discussion since no specific core dimensions are discussed.

General Description of the BRA

The BRA is basically a self-saturating magnetic amplifier using a single magnetic circuit. In contradiction to conventional magnetic amplifiers, it provides controlled output on both positive and negative power pulses.

Consider the BRA circuit in Fig. 2(A). Voltage wave shapes are shown in Fig. 2(B) with circled numbers. These numbers relate time and are referred to as t_1, t_2, t_3, etc., meaning the instant of time at the circled numbers 1, 2, 3, etc. Flux resets between t_3 and t_4; see Fig. 2(C). For the purpose of this paper, flux reset is defined as setting up the flux level in the reactor core in the unsaturated region. The signal current, or bias, determines the flux change $\Delta\Phi$ during flux reset. The flux is almost reset by t_4. A small flux change occurs between t_4 and t_1 but is negligible compared to $\Delta\Phi$ between t_3 and t_4.

Controlled output e_L Fig. 2(B), is produced between t_1 and t_3. The gate winding voltage e_g and the load divide the power pulse e. The portion of e taken by e_g is determined by the flux change between t_1 and t_3 which is equal to the flux change during flux reset. A transfer characteristic is produced similar to that in Fig. 2(*D*). Point *a* in Figs. 2(*C*) and 2(*D*) marks I_{L+} and flux level after flux reset to correspond with e_L of Fig. 2(B).

If the rectifier bias e_b, Fig. 2(A), were omitted, flux could not reset between t_3 and t_4. The equivalent circuit is shown in Fig. 2(E). A very small flux change induces current i_L which produces ampere-turns i_LN_g sufficient to almost equal the control i_cN_c. i_LN_g oppose i_cN_c and almost cancel the effect of the control. This results in negligible flux change (except in very small reactors) between t_3 and t_4. The rectifier bias e_b permits flux reset between t_3 and t_4. The equivalent circuit, Fig. 2(F), has e_b to hold the rectifier blocking. No current, assuming a perfect rectifier, can flow through N_g as long as e_g is less than e_b. The biased rectifier permits flux reset to any level between positive and negative saturation, as desired by the control. Flux reset between t_3 and t_4 leaves the negative power pulse t_4 to t_6 free for other use. Flux can also reset between t_6 and t_1 in the same manner as between t_3 and t_4.

Push-Pull Output with One Core

The BRA provides push-pull output with one core and one gate winding. A typical circuit is shown in Fig. 3(A). The circuit is similar to a half-wave magnetic amplifier, except rectifier *BR2* is added in parallel and oppositely poled to rectifier *BR1*. Each rectifier is biased with a voltage source such as a battery. *BR2* permits negative load current I_{L-} to flow during the negative power pulse. Since flux reset occurs between power pulses, I_{L-} is controlled by the signal similar to the load current during the positive power pulse I_{L+}. I_{L-} flows through the gate winding opposite to I_{L+} and is controlled oppositely. A positive signal increases

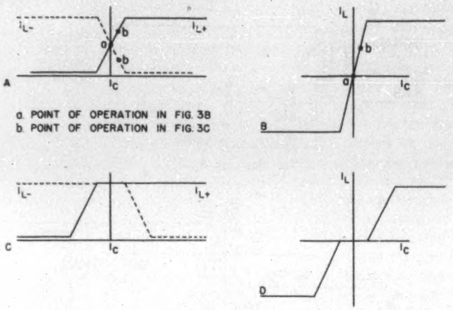


Fig. 4. Transfer characteristics of Fig. 3

A and B—Rectifier and load currents
C and D—Same without a-c bias

I_{L+} and decreases I_{L-}. By subtracting I_{L-} from I_{L+} push-pull output is obtained. Although the pulses providing I_{L+} and I_{L-} are separated in time, they can be subtracted by filtering, such as in a d-c meter, magnetic discriminator, etc. Typical transfer characteristics of I_L, I_{L+}, and I_{L-} are shown in Figs. 4(A) and 4(B).

FULL WAVE WITH ONE CORE

The BRA provides full-wave single-sided output (not push-pull) with one core and one gate winding. A circuit with its transfer characteristic and load-voltage wave shape is shown in Fig. 5(A). Basically, the pulse power supply is rectified full wave, and not filtered. The rectifiers and bias e_b are used to permit flux reset between power pulses. It provides identical unidirectional output on both positive and negative power pulses. This provides twice as much controlled output as the same core could produce in half-wave operation.

CAPACITIVE LOAD

In addition to filtering the output, a capacitive load can supply the rectifier bias, Fig. 5(BI), and increase the output power. Compare Fig. 5(BIV) to Fig. 5(AIV) and the test results in Fig. 6. The capacitor holds voltage between pulses which serves in place of the rectifier bias E_b. Properly applied capacitive loads can also serve in place of rectifier bias for the push-pull circuits. In many push-pull applications, the capacitors are placed across the "dummy leg" resistors. Three examples are shown in Figs. 7(B), (C) and (D) [c replaces of E_b of Fig. 7 (A)].

LOW DRIFT

The BRA reduces drift below the level of conventional magnetic amplifiers. The single core for push-pull eliminates core balance problems. Most of the BRA circuits, for instance, Fig. 3(A), reduce the back voltage across the rectifiers, reducing rectifier leakage and drift. The BRA does not have as low a drift as its partner the pulse relaxation amplifier.[1] Some small selenium rectifiers appear to drift more than the standard size rectifiers, but indications are that silicon and germanium rectifiers promise excellent results in BRA circuits in many applications.

List of Symbols

$BR1$ = biased rectifier, conducts when e is positive
$BR2$ = biased rectifier, conducts when e is negative
C = capacitor used to provide rectifier bias
E = supply voltage, volts average between t_1 and t_4 or t_4 and t_1
e = supply voltage, volts instantaneous
E_b = bias voltage, volts average between power pulses or direct current
e_b = bias voltage, volts instantaneous
E_g = gate winding, volts average between t_1 and t_4 or t_4 and t_1
e_g = gate winding, volts instantaneous
E_L = load voltage, volts average between t_1 and t_4 or t_4 and t_1
e_L = load voltage, volts instantaneous
E_m = peak supply voltage, volts maximum
E_r = rectifier voltage, volts average between t_1 and t_4 or t_4 and t_1
e_r = rectifier voltage, volts instantaneous
f = frequency of supply voltage, cycle per second
I_c = control current, amperes d-c
i_c = control current, amperes instantaneous
I_L = load current, amperes, average over full cycle of e (t_1 to t_1)
I_{L+} = positive load current, amperes, average resulting from e positive

Vol
Pt.
195

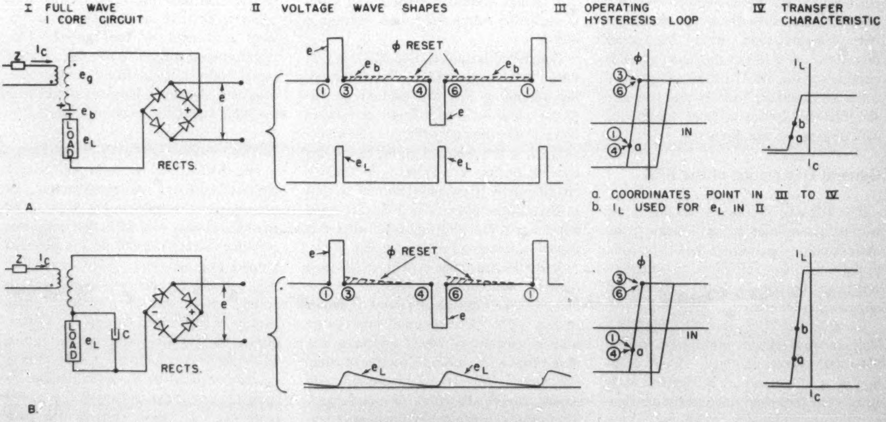


Fig. 5. Full-wave BRA with rectifier bias

A—Fixed voltage source B—Capacitor

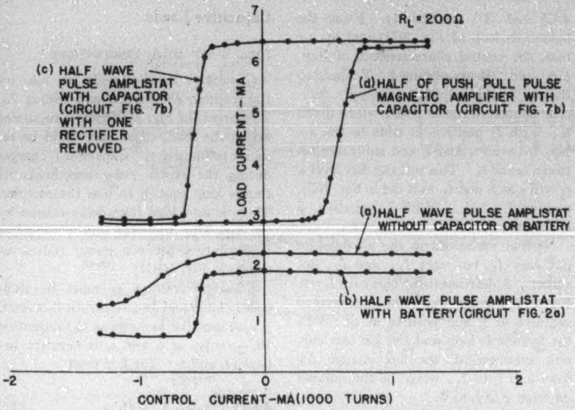


Fig. 6. Transfer characteristics of a pulse magnetic amplifier

I_{L-} = negative load current, amperes, average resulting from e negative
i_L = load current, amperes, instantaneous
mmf = magnetomotive force, ampere-turns per inch
N_c = number of turns in the control winding
N_g = number of turns in the gate winding (power winding)
t = time, instantaneous as specified by numbers in Figs. 2(B) and 5(AII); (time instances are referred to as t_1, t_2, t_3, t_4, etc., indicating instant of time at circled numbers 1, 2, 3, 4, etc.
t_{s+} = time positive saturation is reached, time instantaneous
t_{s-} = time negative saturation is reached, time instantaneous
z = impedance, ohms
Φ_{s+} = flux level at positive saturation, maxwells $\times 10^{-8}$
Φ_{s-} = flux level at negative saturation, maxwells $\times 10^{-8}$
$\Delta\Phi$ = fluxlevel change between specified limits, maxwells $\times 10^{-8}$
φ = flux level, maxwells $\times 10^{-8}$
$d\varphi$ = flux level change, maxwells $\times 10^{-8}$ instantaneous

Theory of Operation

Basic Principle

The half-wave circuit of Fig. 2(A) illustrates the general principle of the BRA. The power supply pulse E has sufficient time voltage to drive the reactor flux from negative to positive saturation between t_1 and t_3. The flux resets[3] mostly between t_3 and t_4. The gate circuit voltages are

$$e = e_g + e_r + e_b + e_L \tag{1}$$

Between t_1 and t_3, normally $(e_r + e_b) << e$, and

$$e_L \cong e - e_g = e - N_g \frac{d\varphi}{dt} \tag{2}$$

At $t = t_1$, $e \cong e_g$ and remains constant until φ reaches Φ_{s+}. Then the gate reactance becomes zero and $e \cong e_L$. The load current

$$I_{L+} \propto \int_{t_1}^{t_4} e_L dt \cong \int_{t_1}^{t_3} edt - N_g \frac{\Delta\Phi}{(t_3 - t_1)} \tag{3}$$

where $\Delta\Phi$ is the flux swing between t_3 and t_3. E.g., with I_c at point a of Fig. 2(D), the flux change $\Delta\Phi$ is equal to the saturation flux Φ_{s+}, $E_g \cong E_L$, and both add up to the supply voltage E.

$\Delta\Phi$ in equation 3 is the same as $\Delta\Phi$ during flux reset. In the circuit of Fig. 2(A) some change of flux occurs between t_4 and t_1 but this change of flux is negligible. At t_3, $\varphi = \Phi_{s+}$ and φ proceeds to drop. The gate voltage $e_g = N_g\ (d\varphi/dt)$ and is negative at the dot in Fig. 2(A). In equation 1, e_L and e are zero, leaving $e_g = e_b - e_r$, assuming complete rectifier blocking. The flux resets between t_3 and t_4

$$\Delta\Phi \propto \int_{t_3}^{t_4} e_g dt = \int_{t_3}^{t_4} e_b dt - \int_{t_3}^{t_4} e_r dt \tag{4}$$

With $e_b = 0$ between t_3 and t_4, the rectifier would conduct, making $e_r \cong 0$, and flux will not reset, as shown n Fig. 2(E). With a finite voltage e_b, the equivalent becomes Fig. 2(F). As long as $e_b \geqq e_g$, the rectifier prevents current flow e_L.

The flux swing, $\Delta\Phi$ in equation 4, depends upon the control current I_c. At $I_c = 0$, $\Delta\Phi = 0$, and the load current I_{L+} is maximum. As I_c increases negative, the flux change remains zero until the control current I_c reaches point b in Fig. 2(D), because of residual in the core; see point b in Fig. 2(C). Then the flux change $\Delta\Phi$ varies with I_c until I_c reaches point c. With I_c at point c, flux drops from Φ_{s+} to Φ_{s-}, the maximum flux change $\Delta\Phi$. As the control current I_c controls the flux change $\Delta\Phi$, during flux reset $\Delta\Phi$ controls the load current I_{L+}, as shown in equation 3.

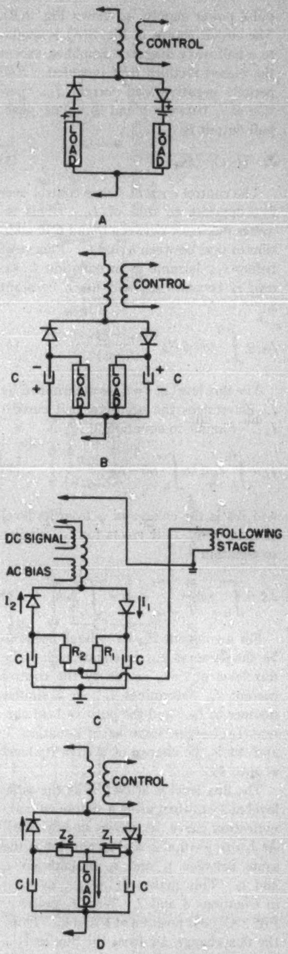


Fig. 7. Four push-pull BRA circuits

Push-Pull BRA with One Core

The push-pull BRA, Fig. 3(A), has one core and one gate winding. It uses a pulse power supply, as shown Fig. 2(B). The circuit shown in Fig. 3(A) is similar to a half-wave magnetic amplifier, except the biased rectifier *BR2* is added. *BR2* permits negative load current I_{L-} (portion of I_L between t_4 and t_6). The push-pull output is

$$I_L=(I_{L+})-(I_{L-}) \tag{5}$$

The control current I_c has control over I_{L+} opposite to that of I_{L-}. This requires flux reset between t_6 and t_1 in addition to that between t_3 and t_4. Flux reset before I_{L-} is same as in equation 4, except e_b becomes e_{b1}, and for a constant e_{b1}

$$E_{b1} \geqq \int_{t_3}^{t_4} e_g dt \leqq N_g \frac{(\Phi_{s+})-(\Phi_{s-})}{(t_4-t_3)} \tag{6}$$

The flux level at $t=t_4$, as determined by I_c, determines the negative load current I_{L-}. Similar to equation 3, I_{L-} is

$$I_L \propto \int_{t_4}^{t_1} e_L dt \cong \int_{t_4}^{t_6} edt - N_g \frac{\Delta\Phi}{(t_6-t_4)} \tag{7}$$

and $\Delta\Phi$ is the change of φ from its level at t_4 to Φ_{s-}. Flux resets for I_{L+} similar to equation 4

$$\Delta\Phi \propto \int_{t_6}^{t_1} e_g dt = \int_{t_6}^{t_1} e_{b2} dt - \int_{t_6}^{t_1} e_r dt \tag{8}$$

For a constant E_{b2} its voltage needs to be the same as E_{b1} in equation 6. The flux level at $t=t_1$, as set by the control current I_c, determines I_{L+} in a similar manner to I_{L-}, and the positive load current I_{L+} is the same as in equation 3, and $\Delta\Phi$ is the change of φ from its level at t_1 to Φ_{s+}.

The flux level φ at $t=t_1$ is at the same level at $t=t_4$ when using a simplified magnetization curve, as shown in Fig. 3(B). At $I_c=0$, $\varphi=0$ at t_1 and t_4, and $\Delta\Phi$ is the same between t_1 and t_3 as between t_4 and t_6. This makes $I_{L+}=I_{L-}$, as seen in equations 3 and 7. With I_c positive, Fig. 3(C), φ is positive at t_1 and t_4. Thus, the flux change $\Delta\Phi$ from the flux level φ at t_1 to $\varphi=\Phi_{s+}$ is less than $\Delta\Phi$ from the flux level φ at t_4 to $\varphi=\Phi_{s-}$ and $I_{L+}>I_{L-}$. With I_c negative, the reverse is true: $I_{L+}<I_{L-}$. The characteristics of I_L, I_{L+}, and I_{L-} versus I_c are shown in Fig. 4(A) and (B).

A-C CROSS BIAS

Certain core materials (Deltamax, Supermalloy, silicon steel, etc.) have hysteresis loops similar to that shown in Fig. 3(D). Using a d-c control, the flux would be at a different level at t_1 and at t_4. The control characteristics of Figs. 4(A) and (B) would become those of Figs. 4(C) and (D) respectively. Using the hysteresis loop of Fig. 3(D) and an a-c bias, the control characteristics of Figs. 4(A) and (B) are obtained. With $I_c=0$ the a-c bias is adjusted to supply sufficient ampere-turns so that $\varphi=0$ at t_1 and t_4. With I_c positive, it adds to the a-c bias between t_6 and t_1 and subtracts between t_3 and t_4. This sets the flux level φ positive at t_1 and t_4, as it did in Fig. 3(C), making $I_{L+}>I_{L-}$. With I_c negative, φ is negative at t_1 and t_4 and $I_{L+}<I_{L-}$.

Normal variations in the a-c bias do not vary I_L, but vary I_{L+} and I_{L-} together. A decrease in a-c bias with $I_c=0$, Fig. 3(D), would make the flux level φ negative at t_1 and positive at t_4. With the hysteresis loop and the a-c bias current symmetrical, the flux change $\Delta\Phi$ from φ at t_1 to Φ_{s+} would be the same as $\Delta\Phi$ from φ at t_4 to Φ_{s-}.

$$\int_{t_1}^{t_{s+}} e_g dt = \int_{t_4}^{t_{s-}} e_g dt \tag{9}$$

where t_{s+} and t_{s-} is the time φ reaches positive and negative saturation respectively. An increase in a-c bias would shift φ at t_1 and t_4 oppositely, but $I_{L+}-I_{L-}=0$, with $I_c=0$. The same as bias in any push-pull circuit, if the a-c bias varies enough that φ saturates at t_1 or t_4, there is a decrease in gain as I_c approaches zero, as shown in Fig. 4(D). A normal variation in a-c bias will not disturb the push-pull characteristic. The a-c bias may be sine wave or pulse.

FULL-WAVE BRA WITH ONE CORE

A BRA circuit as shown in Fig. 5(AI) provides full-wave single-sided (not push-pull) output as shown in Fig. 5(AIV). The rectifiers convert the power supply pulses, similar to Fig. 2(B), to unidirectional. This produces the same output amplitude and polarity during the negative power pulse as during the positive pulse with an average load current $I_L=(I_{L+})+(I_{L-})$.

Flux resets between t_6 and t_1, the same as flux resets between t_3 and t_4. Since the load current I_L is unidirectional, $\varphi=\Phi_{s+}$ at t_3 and t_6 and the flux reset is the same as in equation 4. For a constant E_b, similar to equation 6

$$\int_{t_3}^{t_4} e_g dt = \int_{t_6}^{t_1} e_g dt \leqq E_b \geqq \int_{t_1}^{t_3} edt = \int_{t_4}^{t_6} edt \tag{10}$$

The output is controlled the same as in equation 3. Since $I_{L+}=I_{L-}$, the average output $I_L=2I_{L+}$. Using a core with a hysteresis loop, as shown in Fig. 3(D), the transfer characteric I_L versus I_c is similar to that of Fig. 5(AIV).

Capacitive Loads

FULL-WAVE BRA, SINGLE-SIDED

Capacitive loads provide the bias for the rectifier, eliminating and adding E_b. The circuit of Fig. 5(BI) has a capacitor across the load. The load could be resistive or inductive. Capacitor C charges during the power pulse and holds its charge long enough to bias the rectifier. There is no added time delay caused by the capacitor discharge time if capacitor C discharges between power pulses, as shown in Fig. 5(BII).

Capacitor voltage e_L must be sufficiently large and long to permit flux reset. $\int e_L dt$ must be as large as E_b in equation 10. $\varphi=\Phi_{s+}$ at t_3 and e_g is negative between t_3 and t_4. The flux reset

$$\Delta\varphi \propto \int_{t_3}^{t_4} e_g dt \leqq \int_{t_3}^{t_4} e_L dt \tag{11}$$

The controlled output is similar to that in equation 3. $\Delta\Phi$ is maximum when $\int e_L dt$ is minimum, and minimum $\int e_L dt$ must satisfy equation 11. Minimum $\int e_L dt>0$.

$$\max \int_{t_1}^{t_3} e_g dt < \int_{t_1}^{t_3} edt \tag{12}$$

and

$$\min \int_{t_3}^{t_4} e_L dt \geqq \int_{t_4}^{t_3} edt \tag{13}$$

Commonly, I_L minimum is 20 to 40 per cent of I_L maximum. In a small BRA this is not normally harmful. The operation between t_4 and t_1 is the same as that between t_1 and t_4.

The capacitor increases I_L in Fig. 5(BI) over that of Fig. 5(AI), commonly 10 times the output power. A typical comparison is shown in Figs. 6(B) and (C). I_c is the same for the two curves while I_L is increased. This increases the gain as well as the power output.

PUSH-PULL BRA

Capacitive loads in a push-pull BRA can provide rectifier bias. The application of the capacitor is not as direct as in the full-wave single-sided BRA. Basically, any rectifier bias E_b can be replaced by a capacitor and resistor in parallel. This is not usually desirable, because of power loss in the resistor, unless the resistor is part of the load circuit. Capacitors across the load in Fig. 7(B) provide the rectifier bias in place of the batteries in Fig. 7(A), as in Fig. 5 discussed in the foregoing. Fig. 7(B) has two loads; the two loads could be two fields of a motor or two control windings. In high speed magnetic amplifier resistance is

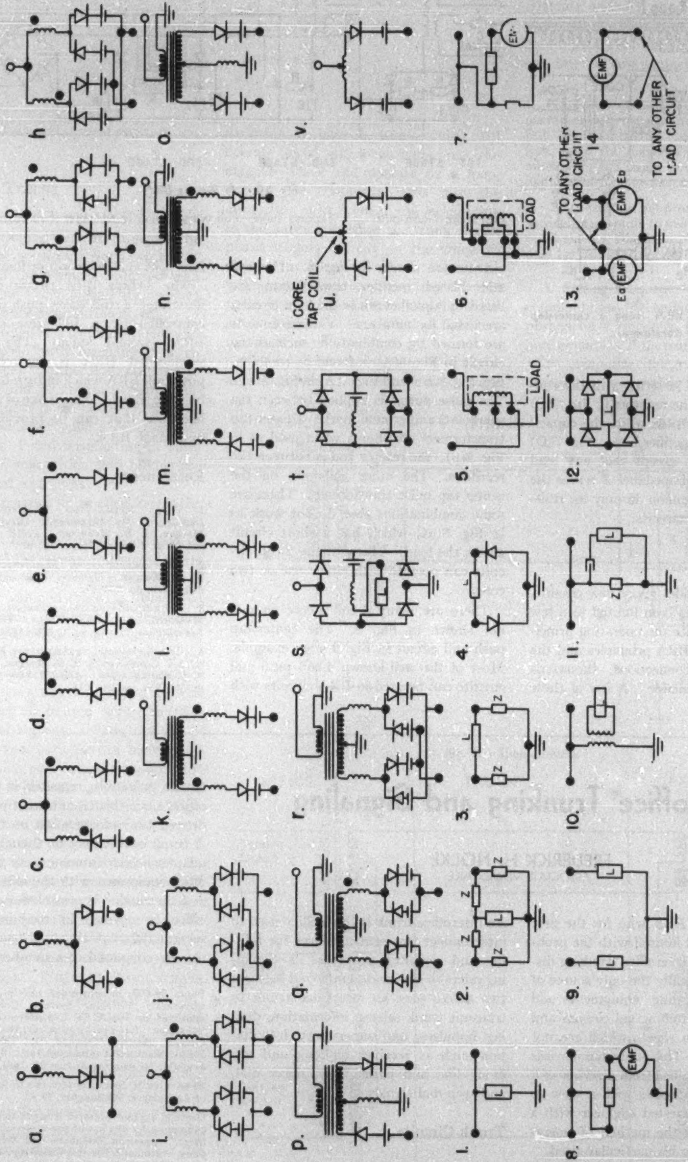

a.
b.
c.
d.
e.
f.
g.
h.
i.
j.
k.
l.
m.
n.
o.
p.
q.
r.
s.
t.
I. CORE
TAP COIL
U.
v.
1.
2.
3.
4.
5.
6.
7.
8.
9.
10.
11.
12.
13.
14
L
Z
LOAD
EMF
Ea
Eb
TO ANY OTHER
LOAD CIRCUIT

Fig. 8. A few BRA circuit combinations

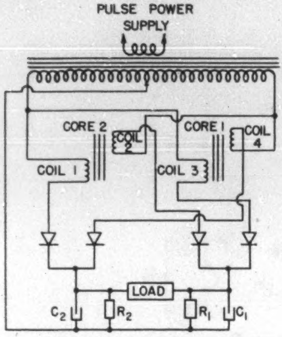


Fig. 9. **Push-pull BRA using a center-tap pulse transformer**

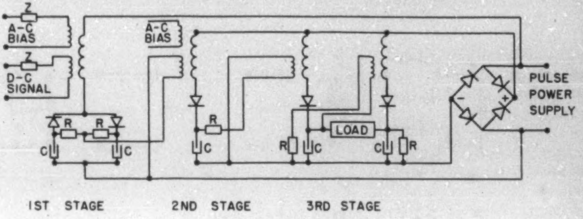


Fig. 10. **A 3-stage BRA**

First stage—Push-pull Second stage—Full wave Third stage—Full-wave push-pull

commonly placed in series with the control windings. The resistors in Fig. 7(C) can serve this purpose while the capacitors provide the rectifier bias. Fig. 7(D) is similar to (C), except that any load may be used and impedance *Z* serves the purpose of the common dummy leg resistors for push-pull circuits.

Ramifications

The BRA provides many new circuits. The discussion has been limited to a few circuits to illustrate the operating principles. Using the BRA principles and the various circuit connections, thousands of circuits are available. A few of these circuits are shown in Fig. 8. The coil and biased rectifier combination are listed by alphabet while the load circuits are listed by numbers. Various circuits are formed by combinations such as the circuit in Fig. 3(A) is found in combination Fig. 8(b1) and Fig. 2(A) in Fig. 8(a1). The pulse power is applied between the symbols ○ and ground mark. Center-tap transformers are shown when used as in Fig. 8(p); the reactor coil is between the rectifiers. The other coils are on the center-tap pulse transformer. There are some combinations that do not work as in Fig. 8(a4) which has a short circuit across the load. Many of the 2-coil circuits can operate with either one or two cores.

There are many 4-coil 2-core circuits not shown in Fig. 8. The center-tap push-pull circuit in Fig. 9 is an example. Most of the well-known 4-coil push-pull circuits can be used in BRA circuits with four coils and two cores or four cores.

The 3-stage BRA circuit of Fig. 10 illustrates a full-wave push-pull circuit controlled by a full-wave single-sided BRA (second stage). The full-wave single-sided BRA is controlled by a push-pull BRA which in turn is controlled by a d-c signal. This is one of many combinations that can be provided with a multistage BRA.

References

1. PULSE RELAXATION AMPLIFIER—1950, A LOW-LEVEL D-C MAGNETIC AMPLIFIER, **R. E. Morgan, J. B. McFerran.** *AIEE Transactions*, vol. 73, pt. I, July 1954, pp. 245–49.

2. THE AMPLISTAT—A MAGNETIC AMPLIFIER, **R. E. Morgan.** *Electrical Engineering*, vol. 68, Aug. 1949, pp. 663–67.

3. FLUX RESETTING CHARACTERISTICS OF SEVERAL MAGNETIC MATERIALS, **Hoobert Huhta.** *AIEE Transactions*, vol. 73, pt. I, May 1954, pp. 111–14.

4. THE EXTENSION OF AMPLISTAT PERFORMANCE BY A-C COMPONENTS, **R. E. Morgan, H. M. Ogle, V. J. Wattenberger.** *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 986–91.

Interoffice Trunking and Signaling

FREDERICK H. NOLKE
ASSOCIATE MEMBER AIEE

EACH ENGINEER who for the first time concerns himself with the problems of telephone interoffice trunking discovers that practically the only source of information regarding equipment and operation comes from actual circuits and circuit description sheets which are not always available. This paper summarizes the information gained from a review of a number of trunk circuits with a view to providing the interested engineer with a basis for choosing the method of operation best suited to his particular need.

For the purpose of this paper the following two definitions are offered: 1. An interoffice trunk is the facility used to interconnect two central offices for talking and signaling purposes. 2. Signaling refers to the means employed between two offices over an interoffice trunk to transmit trunk seizure information, dialing impulses, and supervisory information such as receiver on-hook and off-hook, line and paths busy, delay dial, and stop dial signals.

Trunk Circuits

Any discussion on interoffice trunking and signaling should start with the trunk circuit, or impulse repeater, in the central office, since this circuit must provide the desired type of operation on the trunks. A trunk circuit may be thought of as an adapter which interconnects the central office equipment with the external trunk and thus makes it possible for one central office to operate in conjunction with several different types of trunks. It is usually composed of a number of relays,

---

Paper **54-502**, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 10, 1954; made available for printing August 27, 1954.

FREDERICK H. NOLKE is with the U. S. Department of Agriculture, Washington, D. C.

Grateful acknowledgment is made to the following companies for the use of the material furnished by them: American Telephone and Telegraph Company, Automatic Electric Company, Federal Telephone and Radio Corporation, Leich Electric Company, North Electric Company, and Stromberg-Carlson Company.

a repeating coil, and several resistors and capacitors.

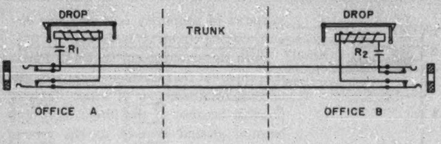


Fig. 1. Ringdown trunk

Ringdown Trunks

A trunk circuit on which it is necessary to furnish ringing current to signal the distant office is called a ringdown trunk. Each end of the trunk is equipped with a-c relays and visual signals for registering the presence of an incoming call.

A 2-way ringdown trunk in its simplest form, equipped with mechanical visual signals (drops) at either end, is shown in Fig. 1. *R*1 and *R*2 are the line drop signals in magneto (local battery) offices *A* and *B* respectively. On a call originating in office *A* and terminating in office *B*, signaling, supervision, and talking are accomplished as follows.

Ringing current is applied from the cord circuit to the trunk at office *A*. This causes relay *R*2 to operate and to "ring down" the mechanical visual signal at office *B*. The operator at office *B* answers the call and connects through to the called line. Automatic on-hook and off-hook supervision is not supplied to either operator, but is accomplished by monitoring the call. Talking battery for the conversation is supplied by the local battery associated with the telephones at either end. When the call is completed and both parties have hung up, the calling party signals his operator with a ring to indicate completion of the call. The operator in office *A* then rings the operator in office *B* over the trunk, actuating a cord supervisory relay, and advises that the call has terminated.

The ringdown trunk is sometimes used between manual common-battery offices. In that instance the signaling may be done by means of "drops" as already discussed or by means of lamps operated by a-c relays. When ringdown trunks are operated in conjunction with common-battery offices the trunk circuit is often equipped to place automatically a spurt of ringing current on the trunk as an indication that the call has been terminated.

Sometimes the situation arises where a 2-way trunk must interconnect a magneto office and an automatic office. That the cost of the facilities in the magneto office may be kept at a minimum, it is often desirable to have a trunk which will operate on a ringdown basis into the magneto office and operate on a loop, simplex, or composite basis into the automatic office. On a call originating in the automatic office the trunk circuit places ringing current on the trunk to signal the magneto office. When the calling party hangs up, the trunk circuit again places ringing current on the trunk as an indication that the call has terminated.

On a call from the magneto office to the automatic office the operator uses a very simple cord circuit which is approximately the equivalent of a telephone set to originate and dial a call into the automatic office. This type of operation has the advantage of being simple and inexpensive but has two distinct disadvantages.

1. The trunk circuits used for this type of operation are usually not convertible to full common-battery operation. Therefore, when the magneto office converts to automatic, the trunk circuits in the automatic office become useless.
2. Since they do not provide the necessary supervisory signals to meet intertoll dialing requirements, ringdown trunks can be used only for trunks to operators.

A ringing frequency of 20 cycles is commonly used when signaling is accomplished on a ringdown basis over short-haul circuits. If the trunk line is equipped with composite telegraph, 135 cycles are used as the signaling frequency so that

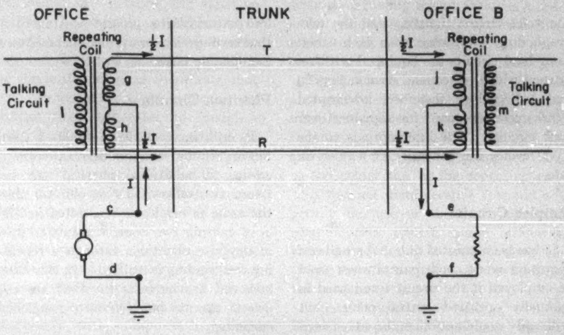


Fig. 2. Simplex trunk

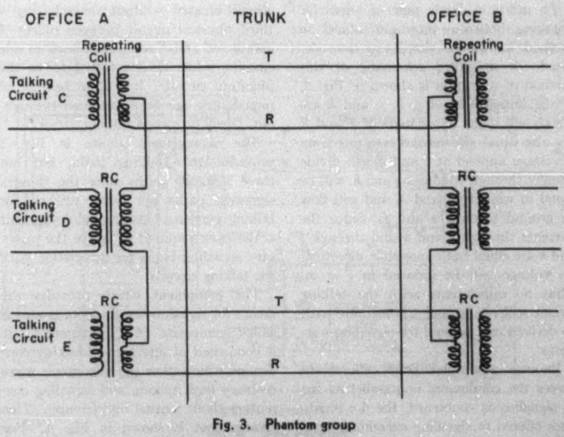


Fig. 3. Phantom group

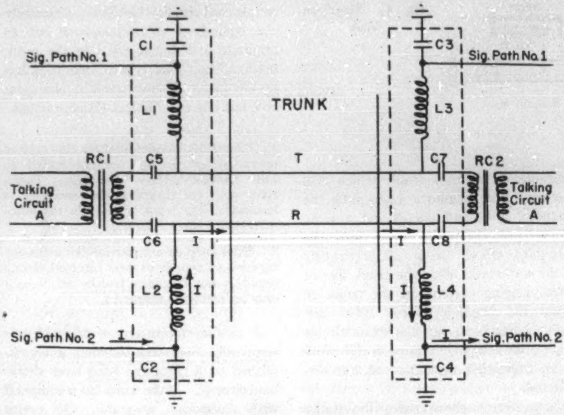


Fig. 4. Composite signaling paths

the voice circuit signaling and the telegraph do not interfere with each other. Long-haul circuits are equipped with repeaters which, of course, do not pass 20 and 135 cycles; therefore, interrupted 1,000 cycles are used for signaling over such circuits. The interruptions to the 1,000 cycles are generally at a 20-cycle rate.

Simplex Circuits

It has been pointed out that ringdown operation over a single pair of wires could be employed if the circuit terminated in manually operated central offices, but ordinarily would not be employed between two automatic central offices.

To utilize a single pair of wires for trunking between automatic offices, a method known as simplexing was devised. A simplified schematic of this method of operation is shown in Fig. 2. If the impedances of *g*, *h*, *j*, and *k* are equal, and if the impedances of *T* and *R* are also equal, the current resulting from a voltage applied at *c* and *d* will divide equally through windings *g* and *h*, will be equal in windings *j* and *k*, and will flow to ground through *e* and *f*. Since the currents through *g* and *h* and through *j* and *k* are equal but opposite in direction, no voltage will be induced in *l* or *m*. Thus no interference with the talking circuit will result, and the metallic path so derived can be used for signaling purposes.

Because a simplex trunk effectively places the conductors in parallel as far as signaling is concerned, the d-c resistance offered to signaling currents is one-fourth (neglecting ground resistance) of that seen by the same trunk operated on a loop-dialing basis.

Phantom Circuits

By utilizing two physical pairs a third talking circuit can be derived without adding an additional physical pair between two offices. To accomplish this, the same principle, as illustrated in Fig. 2, of causing two equal currents to flow in opposite directions through a repeating coil winding is utilized. In this case, however, the currents involved are not direct currents but are voice frequency currents.

In Fig. 3 talking circuit *D* is the new circuit created without necessitating a third physical circuit between offices *A* and *B*. *C* and *E* are now known as side circuits, while *D* is referred to as the phantom circuit. If a close balance of impedances can be maintained throughout, *D* will be a good talking circuit.

The arrangement shown in Fig. 3 provides three talking paths, but not three signaling paths since the simplex signaling paths are being utilized for talking currents. Additional equipment is therefore required to derive the necessary signaling paths for association with the talking circuits.

The equipment which provides the means for signaling on phantom groups is called composite (*CX*) equipment. It is comprised of filters and blocking condensers which allow the side circuit wires to carry both talking and signaling currents without mutual interference. This arrangement is shown in Fig. 4. For purposes of simplicity only one side circuit, talking circuit *A*, is shown.

With the arrangement shown in Fig. 4 it is apparent that two signaling paths are derived from one pair of wires. This is possible because of the filters used and because ground is used as the return path for signaling currents.

The circuit shown in Fig. 4 serves to separate the talking and signaling currents so as to prevent mutual interference. The signaling current *I* contains primarily low-frequency components (below 200 cycles per second). A signaling current, therefore, traveling in accordance with the arrows, sees a high impedance to ground *C*2, a low impedance to the *R* wire through *L*2, a high impedance to *RC*1 (*C*6), a low impedance along the *R* wire, a high impedance to *RC*2 (*C*8), a low impedance *L*4, a high impedance to ground *C*4, and a low impedance into the signaling equipment. A similar set of conditions applies on the *T* wire, yielding two signaling circuits per physical pair.

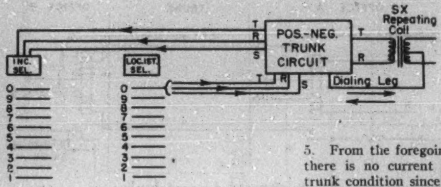
The relatively high-frequency talking currents of from 300 to 3,000 cycles per second see a different set of conditions, with the capacitors *C*5, *C*6, *C*7, and *C*8 representing a very low impedance with respect to that of *L*1, *L*2, *L*3, and *L*4. Thus, the talking currents are restricted to the talking circuit.

Since a phantom group requires two side circuits as shown in Fig. 3, two complete sets of equipment, shown in Fig. 4, are required at each end of a phantom group of trunks. The necessary equipment to provide for the signaling paths at one end of a full phantom group is called a composite, or *CX*, set. When composite equipment is used on a phantom group the trunks are said to be arranged for composite signaling.

It is often desirable from an outside plant point of view to utilize nonquadded cable pairs as trunks, particularly extended-area service trunks. Because phantom groups require a close capacitive balance from each side circuit to ground, a phantom trunk group is usually not operated through more than 3 kilofeet of exchange cable unless measurements are made so that pairs may be chosen with closely matched characteristics. Where trunks are to operate through more than 3 kilofeet of nonquadded cable, simplex operation is preferable.

D-C Signaling

Direct current is usually used for signaling purposes on short-haul circuits. The following signaling methods are typical of the types of d-c signaling most



often encountered:

1. Positive-negative (simplex and composite).
2. Receiving and transmitting (E-M) lead (simplex and composite).
3. Loop.
4. Battery and ground.

Positive-Negative Signaling

Positive-negative signaling has been used for a number of years in the independent telephone industry and has become quite popular. One odd fact in this regard, however, is that there is no term universally applied to this operation. One manufacturer calls it "neutral-differential," another "positive-negative," and other manufacturers seem not to have given it any name.

The method of connecting one 2-way simplex trunk arranged for positive-negative signaling is shown in Fig. 5. This same arrangement would apply for composite (*CX*) operation except that the dialing leg would be connected to one of the signaling paths as shown in Fig. 4 instead of to the center tap of the *SX* repeating coil.

In brief, the principle of operation of positive-negative signaling is as follows:

1. During the idle condition of the trunk, the trunk circuit on each end places ground on the dialing leg through a relay winding. Barring differences in ground potential between the two central offices, there should be no current flow and the relays should not operate.
2. When one end of the trunk is seized by an outgoing call, battery is placed on the dial leg at the originating office, thereby causing the relay in the distant end trunk circuit to operate.
3. Alternately placing battery and then ground on the dial leg at the originating end as a result of dial pulses will cause the relay at the distant end to follow the dial impulses. This relay in turn pulses the central office equipment in the distant office to complete the call.
4. When the called party answers, this fact is passed along back to the originating office by removing the ground on the dialing leg at the distant office and replacing it with battery. Since the originating office also is applying battery to the dialing leg, no current flows in the dialing leg during conversation.

Fig. 5. Positive-negative signaling

5. From the foregoing it is apparent that there is no current flow during the idle trunk condition since ground is applied on both ends of the dialing leg. Likewise, no current will flow during conversation since battery is applied at both ends of the dialing leg. Current flow takes place only during seizure, pulsing, and distant end on-hook supervision on the trunk.

Positive-negative signaling is best suited for use on relatively short trunks and does not provide all of the features that E-M signaling does. A dialing leg resistance of 1,000 ohms is generally the maximum over which these trunks operate. The signaling range of the trunk is affected by such factors as battery voltage, insulation resistance, and earth potentials. These earth potentials may be caused by local interference such as power lines and electrified railroads, or may be the result of geological conditions, magnetic storms, etc. Earth potential differences might prove serious enough to affect the stability of the circuit. A difference of 10 to 15 volts will in most cases prove serious. If the difficulty is owing to 60 cycle alternating current, filters are available which will usually correct the difficulty.

Positive-negative signaling cannot readily be used with carrier equipment. Since a trunk circuit arranged for the E-M type of signaling can be used interchangeably with composite equipment and carrier, it is somewhat more flexible than the positive-negative type of trunk circuit.

In spite of these drawbacks, the positive-negative type of signaling is adequate for most extended area service (EAS) trunks, and is even used in some cases on short-haul toll trunks. Because only relays which are simple in design and relatively inexpensive to manufacture are used, the present cost per trunk is approximately $80 to $100 less than for E-M signaling.

Polar Duplex or E-M Signaling

Polar duplex or E-M signaling may be considered as a refinement of positive-negative signaling. In general, the principles of operation are the same, but the equipment used is somewhat different. With E-M signaling the relays that accomplish the signaling are usually not included in the trunk circuit. Two leads are brought out from the trunk circuit and are connected to the signaling circuit. These receiving and transmitting leads carry the signaling currents. The separate signaling circuit includes a sensitive polarized relay, an artificial line, and an earth potential compensation network.

Fig. 6 shows a composite type of trunk arranged for E-M signaling. This might have been connected as a simplex trunk by removing *C*1, *C*2, *C*3, *C*4, *L*1, and *L*2, and connecting the signaling lead *SL* to the center tap of the repeating coil. This was not done because it would not permit the use of the earth potential compensation circuit, since ordinarily there would not be a metallic path between the offices that could be used for this purpose.

The polarized relay *CX*, shown in Fig. 6, is the heart of the signaling circuit, as will be seen from the following. The *CX* relay has three windings: *W*1, *W*2, and *W*3. *W*3 is used for earth potential compensating purposes and is not directly

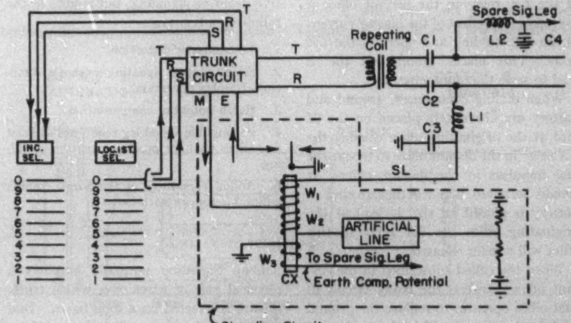


Fig. 6. E-M signaling

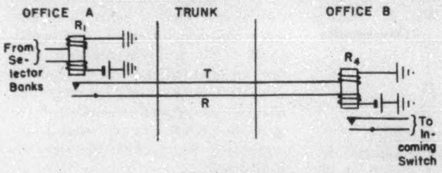


Fig. 7. Basic loop signaling

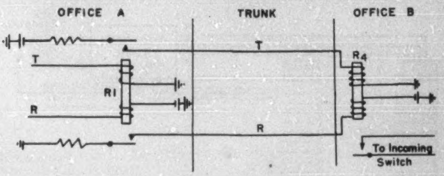


Fig. 8. Battery and ground signaling

involved in the signaling operation. *W*1 and *W*2 have an equal number of turns but the windings are connected differentially. Also included is an artificial line which is adjusted to have approximately the same characteristics as the trunk.

In an idle condition the trunk circuit places ground on the *M* lead and battery through a relay on the *E* lead. In this condition a biasing current flows through *W*2 which holds the *CX* relay in a released position.

When an outgoing call is made, the trunk circuit takes ground off the *M* lead and replaces it with battery. This causes the current through *W*2 to reverse and tends to cause the *CX* relay to operate. However, current also flows through *W*1 and because the voltage applied to *W*2 through the artificial line is higher than ground the current through *W*1 is enough greater than that through *W*2 to hold the relay released. The current through *W*1 flows out over the signaling leg *SL* to the distant office. For the moment, if Fig. 6 is considered to be the distant office, it will be possible to show how an inward call operates in E-M signaling. When battery is applied on the *M* lead at the originating office it is seen that the *CX* relay in the originating office does not operate. However, when this battery is applied to the *W*1 winding of the *CX* relay in the distant office it overcomes the effect of the biasing current in the *W*2 winding and operates the *CX* relay. This places ground on the *E* lead to seize the trunk circuit.

When dialing commences, ground and battery are alternately placed on the *M* lead at the originating office causing the *CX* relay in the distant office to repeat the dial impulses in the distant office. It should be noted that whether ground or battery is placed on the *M* lead at the originating office the *CX* relay in this office will remain released.

When the called subscribed in the distant office answers, the trunk circuit in that office operates to replace the ground with battery on its *M* lead, causing an operating current to flow in its *W*2 winding. Since the two *W*1 windings are now in series to battery at each end of the trunk no current will flow in them. The *CX* relay at the originating office now operates by the current through its *W*2 winding, placing ground on the *E* lead and returning answer supervision to the calling office. The *CX* relay in the distant office remains operated, the operating current in *W*2 taking the place of the previous operating current in *W*1.

The *CX* relay is equipped with a third winding which is used to provide the function of d-c earth potential compensation. Since E-M signaling is designed to operate on a ground return basis, it is apparent that earth potential differences between the two terminals will upset the balance in the system. With the third windings of the *CX* relays at each end of the trunk connected in series over a spare composite leg, a difference in ground potential between the two ends will cause an unbalance current to flow through the third (compensating) winding of each *CX* relay. The resulting effect of this current, therefore, is to compensate for differences in earth potential at the two ends of the signaling leg. This method is effective in compensating for differences in earth potentials up to 50 volts.

As mentioned, the present cost of providing E-M signaling is about $80 to $100 higher per trunk than that for positive-negative signaling, but results in the following advantages:

1. Low-impulse distortion.
2. Simultaneous signaling in both directions (duplex operation).
3. Earth potential compensation.
4. Permissible signal leg resistance at least five times that for positive-negative signaling.
5. Added versatility of the trunk circuit: adaptable for use with carrier.

LOOP SIGNALING

Loop Signaling normally employs a physical pair of wires over which trunk seizure is effected on a loop basis. Dial pulsing is also accomplished on a loop basis, and supervision is effected either on a reverse battery or high-low resistance basis.

Fig. 7 illustrates the basic principles of operation of a loop-dialing trunk. For purposes of simplicity the talking circuit has been omitted and the equipment arrangement has been shown to illustrate circuit operation only on a call from office *A* to office *B*, as follows.

When relay *R*1 is seized by the preceding switch in office *A*, the trunk loop to office *B* is closed, relay 4 operates and seizes an incoming switch in office *B*.

Dialing at office *A* pulses (opens and closes) relay *R*1. This causes *R*4 to open and close alternately (follow the pulses). The action of *R*4 opens and closes the circuit to the incoming switch at office *B*, thereby passing along the dial pulses to the central office equipment. Supervision may be accomplished on either a reverse-battery or high-low basis as discussed in the following.

BATTERY AND GROUND SIGNALING

The loop dialing scheme as described is limited to approximately 1,000 ohms of loop resistance. Since in many cases a trunk loop resistance may exceed this value, another method of operation known as battery and ground signaling has been developed which extends the loop resistance limit up to 2,000 ohms. In effect, it places the batteries in the offices at each end of the trunk in series. Fig. 8 illustrates the basic principles of operation of a battery and ground trunk. For purposes of simplicity, the talking circuit has been omitted and the equipment arrangement has been shown to illustrate circuit operation only on a call from office *A* to office *B*, as follows.

When relay *R*1 is seized by the preceding switch in office *A*, battery is connected to the trunk loop, thus closing the loop through relay *R*4 and battery at office *B*. As can be seen, the battery connections to the trunk are opposite at the two offices, thereby effectively putting the two batteries in series. *R*4 operates and seizes an incoming switch in office *B*.

Dialing at office *A* pulses (opens and closes) relay *R*1. This causes *R*4 to

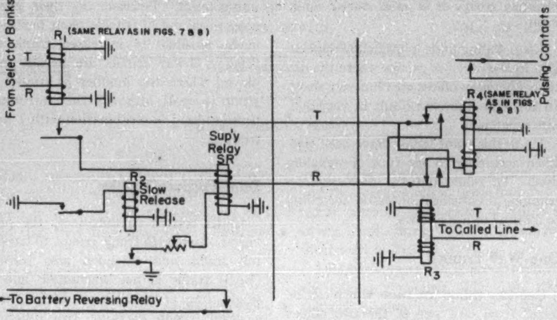

From Selector Banks
R1 (SAME RELAY AS IN FIGS. 7 & 8)
T
R
Sup'y Relay SR
R2 Slow Release
To Battery Reversing Relay
T
R
Pulsing Contacts
R4 (SAME RELAY AS IN FIGS. 7 & 8)
T
To Called Line
R
R3

Fig. 9. Loop signaling, reverse battery supervision

open and close alternately (following the pulses). The action of *R*4 opens and closes the circuit to the incoming switch at office *B*, thereby passing along the dial pulses.

Supervision (the passing back from office *B* to office *A* of on-hook and off-hook conditions on the call) may be accomplished on either a reverse-battery or high-low basis, as discussed in the following.

Reverse Battery Supervision

Reverse battery supervision, as the name implies, is supervision provided by effecting a reversal of battery potential on the trunk. Fig. 9 is the same as Fig. 7 except that it covers answer supervision from office *B* to office *A* in addition to seizure and pulsing.

After the seizure and pulsing relay *R*1 has operated and closed the trunk loop, a current flows through one winding of relay *SR* owing to the battery voltage on the trunk at office *B*. An equal and opposite current flows through the other winding of relay *SR* because the battery loop is closed by the operation of relay *R*1. Since the effects of these currents cancel each other, relay *SR* does not operate.

When the called party at office *B* answers, a loop to battery through relay *R*3 is closed; this relay operates and reverses battery on the trunk at office *B*, which in turn reverses the direction of current through one winding of relay *SR* at office *A*. Since the currents through the windings of relay *SR* are now both in the same direction, the relay operates and passes answer supervision back to the calling link.

When the called party hangs up, battery is removed from relay *R*3 at office *B*, the relay releases and restores battery on the trunk to its original polarity. This reverses the direction of current through one winding of relay *SR* causing it to release and pass on-hook supervision back to the equipment in office *A*.

High-Low Supervision

Supervision is sometimes accomplished on trunks from an automatic to a manual office by effecting a change in resistance of a bridge across the trunk and thereby passing trunk seizure, on-hook, and off-hook information on the call to the distant office. This is called high-low supervision, and is not used on 2-way trunks between automatic offices.

The essential elements of a high-low resistance trunk are shown in Fig. 10. The circuit features at the automatic office end of the trunk only have been shown. The following explanation, however, provides the pertinent information in regard to the operation of this trunk.

With the trunk in an idle condition, it can be seen that battery is placed on the trunk through 280 ohms. A high-resistance relay *M* is connected across the line at office *B*, and the resultant current through this relay is sufficient to maintain it in an operated condition. This maintains the trunk lamp circuit at office *B* in an open condition and furnishes idle (dark-lamp) trunk supervision.

In the case of an outgoing call from office *A*, when the trunk is seized at office *A*, relay *A* operates and relay *G* operates momentarily, opening the circuit to the trunk, thus releasing relay *M*. When relay *G* returns to its original position, the circuit again applies battery to the trunk but this time through 2,280 ohms. This resistance is so high that insufficient current flows through relay *M* at office *B* to reoperate it. This latter action causes the trunk-calling lamp circuit to close, lighting the trunk-calling lamp.

When the operator at office *B* answers the call, the trunk-calling lamp is automatically extinguished. The supervisory lamp associated with the answer cord of the operator's cord circuit also provides dark-lamp supervision at this time.

The operator at office *B* then signals the called subscriber in the customary manner and receives dark-lamp (off-hook) supervision on her calling cord when the subscriber answers. In addition, office *B* places a low-resistance bridge across the trunk, operating relay *C* at office *A* which causes answer supervision to be returned to the calling link.

When the call terminates and the calling party hangs up, relay *A* in office *A* releases, reducing the bridge across the line to 280 ohms. The resulting increase in current causes the answering cord supervisory light in the operator's cord circuit to light up, thereby providing bright-lamp (on-hook) supervision on the call. This action has no effect on the trunk-calling lamp at office *B* and the lamp remains dark. The operator at office *B* then takes down the connection to the called party's line.

In the case of an outgoing call from office *B*, when the trunk is seized by the operator at that office, a low resistance is

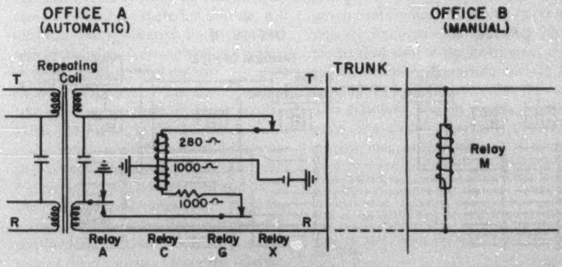


Fig. 10. High-low supervision

placed across the trunk. This operates relay *C* in office *A* which causes an incoming switch in office *A* to be seized.

The impulse springs of the operator's dial (not shown in Fig. 10) at office *B* are in series with the trunk circuit, and pulses from the dial cause the line to be opened and closed alternately. Relay *C* (pulsing relay) at office *A* follows these impulses and passes them to the incoming switch.

When the called subscriber in office *A* answers, battery reversal is provided by the connector. This causes relay *A* to operate, placing 2,280 ohms on the trunk. This latter action causes the calling-cord supervisory lamp to be extinguished, thereby providing dark-lamp (on-hook) supervision on the call. When the called party hangs up, battery reversal is again provided by the connector releasing relay *A*. This again places battery through 280 ohms on the trunk at office *A*. The d-c current on the trunk to office *B* is thereby increased, the supervisory relay operates, and causes the calling-cord supervisory lamp to light up and provide bright-lamp (on-hook) supervision on the call. When the operator at office *B* takes down the connection, the equipment at both ends of the trunks is restored to its normal (idle) condition.

High-low supervision has been largely replaced with reverse battery supervision and is now used principally where additions to existing trunk groups are being made. Although this discussion has been restricted to supervision on loop dial trunks, high-low supervision has also been used to a limited extent on positive-negative type *SX* and *CX* trunks, and sometimes in combination with reverse battery supervision on loop dial trunks.

Loop-dialing trunks have certain restrictions on their use.

1. With the trend toward smaller gauge cables the loop resistance limit is reached sooner.
2. The trunk circuits are not as versatile as the E-M type since they cannot be used on phantom groups or on most carrier equipment.

Loop dialing finds its principal application in 1-way trunk groups where the distances between offices are relatively short. In this case, they can result in very definite economies.

Up to this point trunks have been classified according to the type of signaling used. To round out the discussion on trunking a classification made according to method of use follows.

One-Way Trunks

One-way trunks are those which can be seized from only one of the two offices they interconnect. In general, these trunks are provided on a loop-signaling basis, but simplex or composite operation may also be obtained.

Since 1-way trunks are less economical with regard to outside plant facilities than 2-way trunks, their use is rather limited. Under the following conditions the use of 1-way trunks may be warranted:

1. When 1-way traffic only is involved.
2. When there is a marked difference in the amount of traffic in the two directions.
3. Where the volume of traffic to be handled is high, requiring approximately 25 2-way trunks.

One-way trunk circuits are less costly than 2-way trunk circuits and often no trunk circuit at all is required on the incoming end of a 1-way trunk. Further, an incoming selector is not required on the outgoing end of a 1-way trunk. This saving in central office equipment must be balanced against the added investment required in outside plant, in order to determine whether the use of 1-way trunks is economically warranted.

Two-Way Trunks

Two-way trunks are those which can be seized from either of the two offices they interconnect. Because of their more economical use of outside plant facilities, made possible by increased trunk efficiency, 2-way trunks are usually employed where the number of trunks in group is small, although they are sometimes utilized in combination with 1-way trunks.

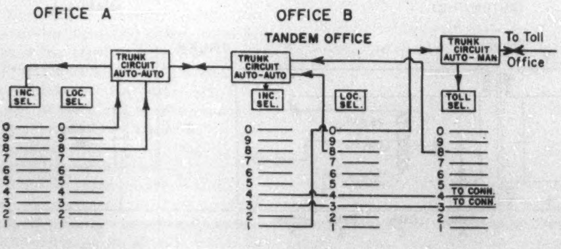


Fig. 11. Tandem operation

Dual-Purpose Trunks

Under certain circumstances the provision of separate trunk groups to handle toll traffic and extended area service (EAS) traffic is not warranted, and a single group of trunks is provided. Such a trunk group carrying two different types of traffic is called a dual-purpose (sometimes called dual-function or dual-use) trunk group.

Fig. 11 illustrates a typical application of dual-purpose trunk use in switch-type equipment. In this instance the trunk group between offices *A* and *B* carries both toll and EAS. Since it is undesirable for the subscriber at office *A* to have to dial other than the single digit *O* to reach the toll operator, appropriate provisions must be made for such a call to "tandem through" an intervening office (office *B*). This is accomplished through the use of a special type of trunk circuit at each end of the trunk group. When (Fig. 11) the trunk group between offices *A* and *B* is accessed from level 8 in office *A* it operates in a very straight-forward manner. The call lands on the incoming selector in office *B* and when an additional digit is dialed it operates the incoming selector to whatever level is dialed.

However, if a subscriber in office *A* wants to make a call to the toll operator, the operation is somewhat different. It is desirable for several reasons for the subscriber to dial only the single digit *O* to reach the operator. To permit this, the trunk circuit in office *A* is of such a type that when it is accessed from level *O* (as opposed to level 8) it will generate an impulse automatically. This automatically generated impulse will cause the incoming selector in office *B* to rise to, and hunt over, the first level. If the first level on the incoming selector in office *B* is multiplied to the *O* level on the local selector the call will be routed automatically through office *B* to the toll center.

On a call from office *B* to office *A* it becomes necessary for the equipment to be able to differentiate between a tool and an EAS call. This is necessary because the toll operator should have access to verification facilities in office *A*

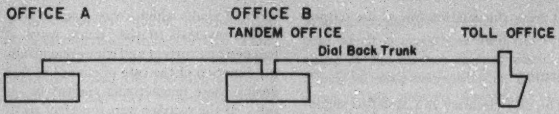


Fig. 12. Dial-back operation

and a subscriber in office *B* should not. In addition, some manufacturers differentiate between the two types of calls so that they may return busy flash and tone on a toll call and busy tone only on an EAS call.

This distinction between inward toll and EAS calls is made, at least by some manufacturers, by generating a marking impulse in the trunk circuit at office *B* on toll calls and no impulse on EAS calls. This again is done on the basis of whether the trunk circuit is accessed from the toll selector or the local selector in office *B*. If a local subscriber in office *B* dials the verification number in office *A* he receives busy tone.

Dial-Back Trunks

In areas served by several community dial offices which reach the operator through a tandem office, and where there is a toll charge between offices, it is sometimes desirable to use dial-back operator office trunks between the tandem office and the operator office, instead of providing an extra (second) group of trunks between these offices. Such a situation is illustrated in Fig. 12.

Dial-back operation is accomplished in the following manner (Fig. 12): On a toll call originating at office *A* and destined for office *B*, the call is tandemed through office *B* to the toll office. The toll operator answers with an answering cord. When she has determined that the call is destined for office *B*, she dials the called subscriber's number back over the same trunk. Upon answering, the called subscriber's line is automatically connected to the trunk. The operator may remain in on the connection, and time the call, or she may withdraw her connection to the trunk, as circumstances warrant. In some types of dial-back equipment the trunk between office *B* and the toll office drops out of the connection when the operator releases. In other types of dial-back equipment this trunk cannot be released until the call is completed. This type of trunk may be desirable where excessive back-haul would otherwise be involved.

In general, trunks of this type have disadvantages from an operating viewpoint for use with an operator office employing a large toll board, because the method of operation is different from that on regular operator office trunks.

Conclusion

This discussion does not purport to cover all phases of interoffice trunking since there are a number of variations in trunk circuits to handle individual situations, a fact which has caused many manufacturing and operating complications and since every attempt toward standardization of trunk circuits always meets with the problem of tying offices in with a nonstandard type of equipment. It is the author's opinion that the independent telephone industry generally would benefit from the adoption of certain standards with regard to trunks. Although this paper has been written to give a brief explanation of the method of operation of the various types of trunks, it is hoped that it will serve also to encourage some thinking toward the objective of standardization.

Single-Ended Saturable Reactor Circuit with Quiescent Current Compensation

R. J. RADUS
NONMEMBER AIEE

Synopsis: This paper has been prepared to introduce a single-ended saturable reactor circuit which exhibits a transfer characteristic for which the d-c output is essentially a linear function of the d-c input for the entire design range of d-c inputs. This "to-zero" operation is shown to be the result of a unique combination of nonlinear characteristics of conventional devices, transformer and rectifier, which eliminates the magnetizing component of the reactor current from the d-c output. The new circuit is presented and a qualitative description of the circuit operation is given.

THE basic component of the circuitry presented in this paper is the conventional series-connected saturable reactor with relatively low control circuit impedance.[1] A brief description of the operation of this basic circuit, as concerns a typical d-c metering scheme, will illustrate the application limitation with which this paper is concerned. A typical metering scheme using the basic circuit is shown in Fig. 1(A). The corresponding transfer characteristic is given in Fig. 1(B). In Fig. 1(A) the d-c input, which is the signal received from the d-c system that is being measured, controls the reactor current i_g which is rectified and fed to the indicating instrument *M*. Thus, the reading on the indicating instrument is continuously representative of the magnitude of the d-c input. Since the relationship between the instrument reading and the d-c input is given by the transfer characteristic, the instrument scale must be calibrated to fit this particular transfer characteristic. An interpretation of the transfer characteristic, as concerns the scale of a linear movement instrument, will show that the scale is linear over the range of d-c inputs for which the transfer characteristic is linear and that the instrument scale is nonlinear over the range of d-c inputs for which the transfer characteristic is nonlinear. In addition, the instrument indication which represents zero d-c input will not coincide with the mechanical zero of the instrument owing to the reactor magnetizing current that flows in the instrument at zero d-c input. The resultant loss in usable length of scale, or stated another way, the ratio of reactor magnetizing current to nominal

---

Paper 54-393, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11-15, 1954. Manuscript submitted June 15, 1954; made available for printing August 25, 1954.

R. J. RADUS is with the Westinghouse Electric Corporation, Pittsburgh, Pa.

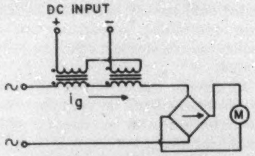


Fig. 1(A). Basic saturable reactor circuit as applied for d-c metering

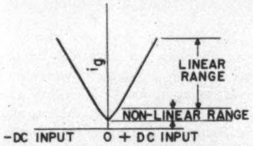


Fig. 1(B). Transfer characteristic of basic saturable reactor circuit as applied for d-c metering

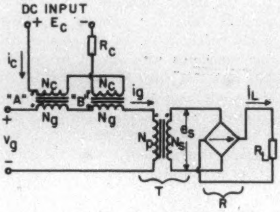


Fig. 2. Single-ended saturable reactor circuit for full-range linear operation

or full-scale reactor current, would be roughly 2 per cent to 10 per cent, depending upon the reactor design for a particular application. This quiescent output and the associated deviation from the inherent linear performance, while not being critically restrictive in the typical d-c metering scheme of Fig. 1(A), are certainly undesirable from a practical application standpoint since they require the use of a nonstandard instrument scale that must be calibrated to fit the given transfer characteristic. It is apparent now that any method which eliminates or minimizes the effect of the magnetizing current without destroying the inherent linearity of saturable reactor operation would be a considerable improvement.

There have been a number of methods devised for eliminating or minimizing the quiescent output of a saturable reactor circuitry.[2] The following material is a description of the circuitry presented in this paper and an explanation of the method by which this circuitry eliminates the magnetizing component of reactor current from the relationship of d-c output versus d-c input.

Circuitry

The new circuitry is a modified single-ended saturable reactor circuit. The modification is the addition of a transformer *T* between the series-connected reactors and the output rectifier *R*. The schematic of the circuit is shown in Fig. 2. Typical designs of the circuit components are as follows.

Reactor:

Core material: Hipernik V, 0.002-inch strip
Core size: 1½ inches inside diameter by 2 inches outside diameter by 1/2 inch
Gating winding: 700 turns
Input or control winding: 5,000 turns

Transformer:

Core material: Hipernik V, 0.002-inch strip
Core size: 1 inch inside diameter by 1½ inches outside diameter by 1/2 inch
Primary winding: 130 turns
Secondary winding: 4,460 turns

Rectifier (bridge construction):

Type: selenium
Cell: 0.05 square inch per cell
Cells per leg: five cells in series per leg

This design is capable of supplying a nominal output current of 3.5 milliamperes into a load resistance of 10,000 ohms for an input current of 17 milliamperes d-c. Laboratory testing of this design has shown that the d-c output is a linear function of the d-c input within the accuracy limits of ±0.5 per cent of the nominal output for the full range of d-c inputs. This means that the actual d-c output does not deviate from a truly linear output at any point by more than ±0.0175 milliampere. The operational accuracy is based on full output to conform to standard metering practice.

Operation

The to-zero operation of this circuit is made possible primarily due to the nonlinear characteristics of transformer *T;* however, the nonlinear characteristic of the rectifier *R* also has an important, if somewhat secondary, role; refer to Fig. 2.

The transformer effectively acts to shunt the reactor magnetizing current from the rectifier, hence from the load resistance R_L. The transformer is not entirely effective since the exciting magnetomotive force due to the reactor magnetizing current and the transformer primary turns does produce a flux change in the transformer core. This flux change causes a voltage to be induced in the secondary winding. This induced voltage is a series of peaks which correspond in time to the reversals in the square wave of magnetizing current and whose magnitude is a function of the rate of change of flux during these magnetizing current reversals. If the rectifier were a perfect rectifier, it would conduct at every value of voltage greater than zero voltage and a proportional current would flow in the load resistance. In actual operation, the rectifier, because of its nonlinear characteristic, does not have appreciable conduction at low values of alternating voltage. This inability of the rectifier to conduct appreciably at low values of alternating voltage is commonly called the threshold effect of the rectifier. In the operation of this circuit, the rectifier threshold effect is used to block the voltage which is induced

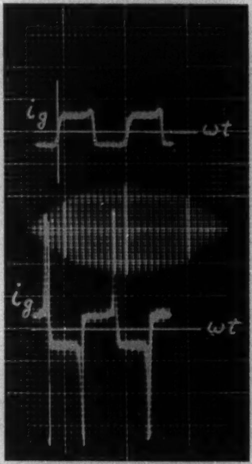


Fig. 3(A) (top). Reactor magnetizing current wave shape

Fig. 3(B) (bottom). Reactor current wave shape showing the conducting pulse

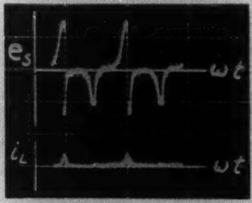


Fig. 4(A) (top). Transformer secondary voltage wave shape at zero d-c input

Fig. 4(B) (bottom). Load current wave shape at zero d-c input in the secondary winding by the action of the reactor magnetizing current. Thus, the shunting action of the transformer nonlinearity and the blocking action of the rectifier nonlinearity are effective in keeping the magnetizing current from flowing in the load resistance and the d-c output will be essentially zero when the d-c input is zero. The wave shape of the magnetizing current is shown in Fig. 3(A). The corresponding induced secondary voltage wave shape and the d-c output current wave shape are shown in Figs. 4(A) and (B) respectively. A comparison of the wave shapes of Figs. 4(A) and (B) shows the effectiveness of the rectifier in blocking the induced secondary voltage for this condition of zero d-c input.

Fig. 3(B) is the wave shape of the reactor current i_g at a finite value of d-c input and shows the conducting pulse in the reactor current. When these currents flow through the transformer primary, the secondary induced voltage becomes greater than the rectifier threshold voltage and the rectifier conducts. At first glance, it would appear that the current which flows in the secondary circuit might be appreciably affected by the nonlinear resistance of the rectifier. This is not the case, however, since the transformer acts essentially as a current transformer with the associated current forcing in the secondary circuit. This current transformer action is the result of having an effectively large resistance in series with the transformer primary, thus allowing the reactor current to be forced through the primary winding unaffected by nonlinear characteristic of the reflected secondary burden. A comparison of Figs. 5(A) and (B) shows evidence of this transformer action. Both figures are plots of the flux level versus primary ampere-turns of the same transformer being excited by a given reactor current of the same general wave shape as that shown in Fig. 3(B). Fig. 5(A) is the plot with the burden connected to the secondary, and Fig. 5(B) is the plot with the secondary open-circuited. Fig. 5(B) shows the flux change that results from the primary current alone. Fig. 5(A) shows that, when both primary and secondary currents are flowing, the flux change is less. It is assumed that this flux change in the core is only that change which is required to produce the voltage that is developed in the secondary burden. This latter type of flux change is characteristic of the operation of current transformers.

The general shape of the plot of Fig. 5(A), transformer core flux versus primary ampere-turns, indicates a form of delayed magnetization in the transformer core which is consistent with the wave shape of the reactor current when the circuit is being operated at a finite d-c input. This delayed magnetization is more clearly shown using individual plots of instantaneous reactor current and instantaneous transformer core flux level. These individual plots, Figs. 6(A) and (B) respectively, have the same time base. A study of these plots shows that the increase in reactor current, Fig. 6(A) from point 1 to point 2, does not cause a corresponding rapid increase in transformer core flux. In Fig. 5(A) this delayed magnetization is shown as the *a* to *b* section in which the rapid increase in reactor current produces only a small change in transformer core flux. Continuing the explanation and again using the individual plot of Fig. 6(B) shows that the core flux continues to change from the level at point 1 for the duration of the reactor current conducting pulse until point 2 is reached. This is the section *b* to *c* on Fig. 5(A). At the end of this conducting pulse in the reactor current point 3 of Fig. 6(A), the reactor current reverses and rises to the magnetizing current level. The corresponding resetting flux change is from point 2 to point 3 in Fig. 6(B). After the resetting period is completed, the next pulse occurs in the reactor current. In Fig. 5(A) the flux resetting period is given as the section *c* to *d*. The bright dot at point *d* indicates that the reactor current and the core flux remain relatively constant for the period π to $(\pi+\alpha)$ of Fig. 6(A). In Fig. 6(B) this is the section point 2 to point 3. At $(\pi+\alpha)$ the reactor current, Fig. 6(A), increases rapidly from point 4 to point 5. The delayed magnetization is effective for this pulse in the reactor current as it was for the previous pulse in the period, α to π, and at the end of the current pulse, point 6 of Fig. 6(A), the flux has changed to point 4 in Fig. 6(B) and point *f* in Fig. 5(A). The core flux is again subjected to the resetting action of the magnetizing current and is then at the flux level given by point 5 of Fig. 6(B) and point *a* of Fig. 5(A). This completes the explanation of the relationship of transformer core flux versus a given reactor current for 1 cycle of the supply frequency. For other reactor currents, i.e., where the reactor current conducting pulses occur at a time either earlier or later in the cycle than $\omega t=\alpha$ and $\omega t=(\pi+\alpha)$ of Fig. 6(A), the plot of the core flux versus reactor current has the same general shape and the preceding explanation is valid.

This explanation will also be substantially true when the operating hysteresis loop of the transformer core is investigated. To convert the plot of Fig. 5(A) to a hysteresis loop, it is necessary to consider the action of the secondary current as well as the primary or reactor current. Fig. 7(A) combines the corresponding secondary current to the plot of Fig. 5(A) and changes the abscissa from primary ampere-turns to primary ampere-turns minus secondary ampere-turns. Applying the previous explanation to Fig. 7(A) shows that the increase in reactor current from 1 to 2 of Fig. 6(A) causes secondary current to flow such that the differential

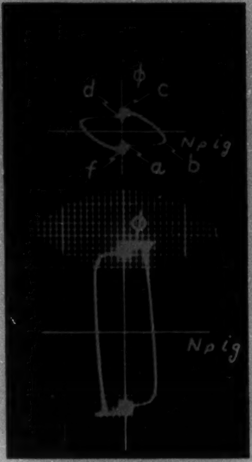


Fig. 5(A) (top). Transformer core flux versus primary magnetomotive force

Fig. 5(B) (bottom). Transformer core flux versus primary magnetomotive force

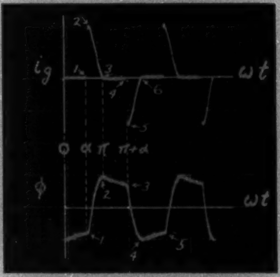


Fig. 6(A) (top). Reactor or primary current versus time

Fig. 6(B) (bottom). Transformer flux versus time

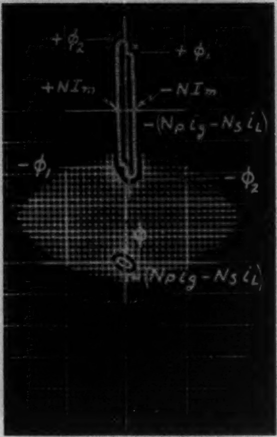


Fig. 7(A) (top). Transformer core hysteresis loop for the postfiring condition

Fig. 7(B) (bottom). Transformer core hysteresis loop for the prefiring condition

ampere-turns are maintained relatively constant at the value of NI_m while the flux changes from $-\phi_1$ to $+\phi_2$. After the flux has reached a maximum at $+\phi_2$, and this occurs essentially at the end of the primary current pulse, it appears that the combination of the phase displacement between the primary and secondary currents and the change in primary current to the magnetizing current level is responsible for the core flux reset from $+\phi_2$ to $+\phi_1$. The resetting continues until the next pulse in the reactor or primary current occurs. The resultant secondary current flow is such that the differential ampere-turns are maintained relatively constant at a value of $-NI_m$ while the flux changes from $+\phi_1$ to $-\phi_2$. The resetting action is again effective and the flux changes from $-\phi_2$ to $-\phi_1$, completing the operation for 1 cycle.

Thus, Fig. 7(A) represents the hysteresis loop of the transformer core when the circuit is operated at a finite d-c input. Fig. 7(B) shows the hysteresis loop of the transformer core when the circuit is operated at zero d-c input. These two different shapes of transformer core hysteresis loops, which typify the nonlinearity that is utilized in the operation of this circuit, result from the two conditions of reactor operation. The prefiring condition is defined as the operation at zero d-c input for which the reactor current is only the magnetizing current of the reactors. The postfiring condition is defined as the operation at finite values of d-c input for which the reactor current includes the magnetizing current and the conducting pulses.

As was stated previously, the to-zero operation of this circuit or the circuit's ability to eliminate effectively the reactor magnetizing current from the d-c output involves the nonlinear characteristic of both the transformer and rectifier. However, if the transformer is designed in accordance with certain predetermined requirements, it will be capable of absorbing the magnetizing current without producing an appreciable secondary voltage, thus minimizing the blocking action which is required of the rectifier. Since the nonlinear characteristic of the transformer is more stable than the nonlinear characteristic of the rectifier, this is a logical division of component importance as concerns the circuit operation for varying environmental conditions. The following analytical method is used to determine the requirements for optimum transformer performance as concerns the to-zero operation of this circuit.

The circuit, shown schematically in Fig. 2, consists of a pair of reactors, A and B, a transformer T, a bridge rectifier R, and a load resistance R_L. The reactors are constructed with high-permeability rectangular-loop core material, such as Hipernik V, a gating winding of N_g turns, and a control winding of N_c turns. The transformer is also constructed with rectangular-loop core material. The transformer has the conventional primary and secondary windings of N_p turns and N_s turns respectively. The reactor gating windings are series-connected and are connected in series with an a-c supply voltage v_g and the transformer primary. The transformer secondary is connected to the a-c terminals of the bridge rectifier and the load resistance is connected to the d-c terminals of the rectifier. The reactor control windings are connected for series opposition to the fundamental frequency of the a-c supply voltage and in series with a resistance R_c and a d-c control or input voltage E_c. The instantaneous currents are defined as follows: i_c in the control windings, i_g in the gating windings and the transformer primary winding, and i_L in the transformer secondary, rectifier, and load resistance. The term (lH) is used to define the instantaneous ampere-turns which are acting on the magnetic circuits and is derived from the expression $H=(0.4\pi NI)/(l)$. The circuit operation will be analyzed first for the prefiring condition and then for the postfiring condition. The circuit equations for the prefiring condition are as follows

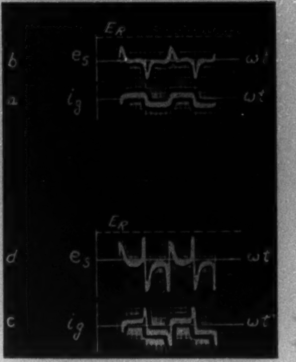


Fig. 8. A—Reactor current for the prefiring condition. B—Secondary voltage for the prefiring condition. C—Reactor current for the postfiring condition. D—Secondary voltage for the postfiring condition

$$(lH)A = i_cN_c + i_gN_g$$

$$(lH)B = -i_cN_c + i_gN_g$$

$$(lH)T = i_gN_p - i_LN_s$$

$$v_g = N_g\phi_A + N_g\phi_B + N_p\phi_T$$

$$E_c = +N_c\phi_A - N_c\phi_B + i_cR_c$$

By definition of the prefiring condition $N_g\phi_A = N_g\phi_B$ and $E_c = 0$. Therefore $i_c = 0$ and $(lH)A = (lH)B = i_gN_g$.

With

$$i_g = (lH)A(1/N_g) = (lH)B(1/N_g)$$

$$(lH)T = (lH)A(N_p/N_g) - i_LN_s$$

Solving for i_L, where

$$i_L = (lH)A(N_p/N_sN_g) - (lH)T/N_s$$

it becomes evident that i_L approaches a zero value as the value of $(lH)T$ approaches the value of $(N_p/N_g)\ (lH)A$. In effect, this means that the magnetizing current of the reactors must be absorbed in the transformer without producing an appreciable flux change. This suggests that the ampere-turns $(lH)T$ which are produced by the reactor current and the primary winding should be equal to or less than the ampere-turns which correspond to the d-c loop width of the transformer core $(lH)T_{\text{d-c}}$. Published data concerning the characteristics of high-permeability rectangular-loop core material, such as Hipernik V, show a relatively small flux change results when the a-c exciting ampere-turns of a core are limited to a value which corresponds to the d-c loop width. For the Hipernik V

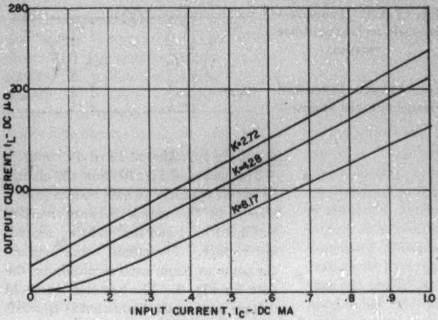


Fig. 9. Experimental transfer characteristics

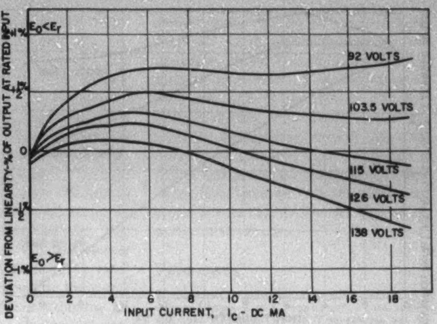


Fig. 11. Experimental results for supply voltage variations

cores that were used in the model that was experimentally tested for this paper, the ratio of a-c loop width at 60 cycles per second (cps) to d-c loop width is roughly 4 to 1. In general, the expression $(lH)T=(N_p/N_g)\ (lH)A$ should be satisfied for $(lH)T=(lH)T_{\text{d-c}}$. Thus $(lH)T_{\text{d-c}}=(N_p/N_g)\ (lH)A_{\text{a-c}}$ is a suggested transformer design prerequisite for minimizing the importance of rectifier blocking.

In the postfiring condition $i_g=(lH)A/N_g-(i_cN_c/N_g)$ and $i_LN_s=-i_c(N_cN_p/N_g)+(lH)A(N_p/N_g)-(lH)T$. This expression shows that the relationship between i_L and i_c will be simply a function of the reactor and transformer turns-ratios if the condition of $(lH)A(N_p/N_g)=(lH)T$ is satisfied. $(lH)A$ is the magnetizing ampere-turns of the reactors as defined in the analysis of the prefiring condition. During the conducting pulse of the postfiring condition, when reactor *B* is saturated, the value of $(lH)A$ is determined by the flux change in reactor *A*. Since, by definition of low control circuit impedance operation, this flux change is relatively small, the resulting rate of change of flux is less than is experienced prior to the saturation of reactor *B*. This then implies that the operating loop width during the conducting period is narrower than the loop width at the supply frequency and suggests that, for the postfiring condition, $(lH)A=(lH)A_{\text{d-c}}$. $(lH)T$ is the magnetizing ampere-turns of the transformer and is defined here again as the difference between the primary and secondary ampere-turns. Fig. 7(A) is a plot of flux level in the core versus the difference in primary and secondary ampere-turns and shows the general shape of the hysteresis loop of the transformer core in the postfiring condition. This shape suggests that in the postfiring condition the effective transformer core hysteresis loop width is roughly the a-c loop width at the supply frequency. As such, it must be greater than the loop width required in the prefiring condition which has been defined as the d-c loop width.

Thus, there are two transformer design prerequisites for consideration.

Prefiring

$$(lH)T_{\text{d-c}}=(N_p/N_g)(lH)A_{\text{a-c}}$$

Postfiring

$$(lH)T_{\text{a-c}}=(N_p/N_g)(lH)A_{\text{d-c}}$$

It is not possible to satisfy these expressions simultaneously, since, in the change from prefiring operation to postfiring operation, the increase in required magnetization of the transformer core is accompanied by a decrease in reactor core magnetization rather than the necessary increase. Because the primary concern of this paper is the effective elimination of quiescent output, the prefiring condition is assumed more important. The resulting loss of reactor current, owing to the required magnetization of the transformer in the postfiring period, will be recognized as a source of error in the ampere-turns relationships of the components. This error can be reduced if the prefiring primary ampere-turns are in-

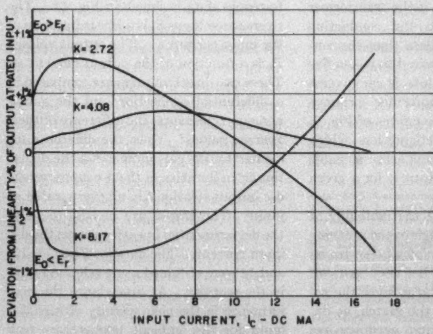


Fig. 10. Experimental performance characteristics

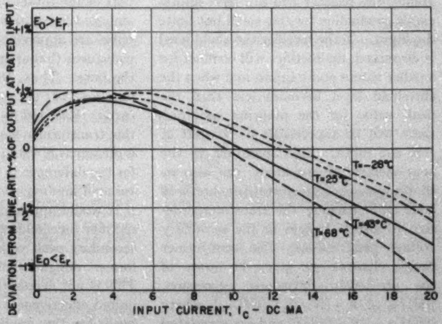


Fig. 12. Experimental results for ambient temperature variations

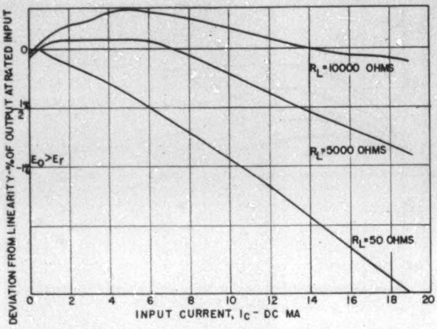


Fig. 13 (left). Experimental results for load resistance variations

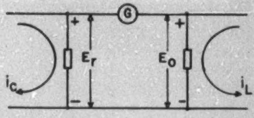


Fig. 14 (right). Differential measuring circuit schematic

creased beyond the d-c loop width and if the rectifier is used to block the resulting secondary induced voltage.

The wave shape of the voltage induced in the transformer secondary for the prefiring condition is peaked owing to the substantially square wave shape of the reactor magnetizing current and the nonrectangular hysteresis loop of the transformer core, Fig. 7(B), at this relatively low level of magnetization. It is this peak value of voltage that must be blocked by the rectifier threshold effect.

The description of the operation for the prefiring condition is shown in Fig. 8. The reactor magnetizing current is shown in Fig. 8(A). The corresponding secondary voltage is shown in Fig. 8(B). The voltage level E_R in Fig. 8(B) represents the threshold voltage of the rectifier. The reactor current wave shape for a finite d-c input is shown in Fig. 8(C). The secondary votage which corresponds to the conducting pulse of reactor current is added to the information of Fig. 8(C) and the total is shown in Fig. 8(D). This latter voltage pulse is also shown to be less than the rectifier threshold voltage. Hence, the rectifier will not have appreciable conduction for this small but finite d-c input. If the rectifier threshold level is decreased, the rectifier will conduct for smaller values of d-c input, and when the threshold level becomes less than the peak value for the prefiring condition there will be appreciable d-c output at zero d-c input. This operation, in the near vicinity of d-c input, can also be shown by using a given rectifier threshold level and changing the transformer design to affect changes in the secondary voltage peak values. The transformer design changes are given in terms of the previously developed expression, $(lH)T_{60\text{ cps}}/K=(N_p/N_0)\ (lH)A_{\text{a-c}}$, where $(lH)T_{60\text{ cps}}/K=(lH)T$, and are correlated to the circuit operation by the curves of Fig. 9. The term $(lH)A_{\text{a-c}}$ is held constant at approximately 2.01 AT. The transformer turns-ratio is held constant to maintain constant loading while the primary turns N_p are changed to obtain values of K equal to 2.72, 4.08, and 8.17. It is seen in Fig. 9 that the best to-zero linearity results when $K=4.08$. With $K=2.72$, the to-zero linearity is destroyed because of the appreciable output at zero d-c input. With $K=8.17$, the to-zero linearity is destroyed because the output remains negligible for an appreciable range of d-c inputs.

Although this apparent correlation between transformer design and circuit operation is consistent with the preceding explanation of the operation, the explanation, while being true, is not complete since it does not point out the effect of the required magnetization of the transformer core. In the previous analysis of the postfiring condition, it was shown that a part of the conducting pulse of the reactor current is used to magnetize the core and is not effective in producing a flux change. For small values of d-c input this effective loss can be an appreciable part of the total pulse if the transformer ampere-turns prior to the conducting pulse are appreciably less than the ampere-turns that are required to magnetize the core. Thus, the loss is an inverse function of the transformer primary turns. Referred to the curves of Fig. 9, this transformer magnetizing loss, which is greater at $K=8.17$, is in part responsible for the difference in output i_L for a given value of input i_c.

It would appear that the matching of rectifier threshold voltage to the prefiring secondary peak voltage of a given transformer design would be very critical. This is not necessarily true since the required effectiveness of the match, as defined by the application accuracy requirements, is determined relative to the postfiring operation at rated d-c input.

The curves of Fig. 10 show the change in circuit operation which results from a change in the match between rectifier threshold level and the prefiring secondary voltage. The circuit conditions are the same as those used in obtaining the data for Fig. 9. The rectifier threshold voltage is maintained constant at approximately 4 volts. The peak values of the prefiring secondary voltages are 1.5 volts, 4 volts, and 5 volts for K values of 8.17, 4.08, and 2.72 respectively. These curves show that the circuit operation does not exceed the accuracy limits of ±1 per cent for the relatively wide range of prefiring secondary voltage.

Performance

This circuit exhibits the performance stability which is characteristic of saturable reactor circuitry since it is essentially a series-connected saturable reactor and uses rectangular-loop core material in the reactors. This performance stability includes the substantial independence of d-c output current from changes in supply voltage and frequency and load circuit resistance.

Laboratory performance testing, which has been conducted on a model according to the components' designs given in the section entitled "Operation," has shown results which should be satisfactory for most d-c metering applications. The results of these tests are given in graphic form in Figs. 11, 12, and 13. The circuitry involved in obtaining these performance data is shown in Fig. 14. The reference voltage E_r is a linear function of the input current i_c. The output voltage E_o is a function of the output current i_L. These two direct voltages are compared in a differential connection and the galvanometer G measures the difference of their average values. Thus, the difference indicated by the galvanometer is the deviation from linearity of the d-c input versus d-c output relationship at every value of input. The graphs are plotted to show the deviation from linearity versus the d-c input current. The data for Figs. 11, 12, and 13 were obtained using selenium cells in the rectifier. At first glance the performance in the near vicinity of zero d-c input for the ambient temperature test data of Fig. 12 might appear unusually stable. However, a qualitative investigation of the components' characteristics shows that a change in rectifier characteristic (e.g., a decrease in threshold voltage for an increase in temperature) is compensated to some degree by the corresponding change in reactor core characteristic (i.e., a decrease in hysteresis loop width for an increase in temperature). The effect of varying ambient temperature on the transformer operation is not readily definable for zero d-c input owing to the relatively low level of transformer core magnetization.

Application

The saturable reactor type of magnetic amplifier or transductor is used as a metering magnetic amplifier because of the inherent linear relationship between the circuit's d-c input and d-c or average value of output. Devices which use the conventional circuitry have found widespread use in isolating the metering circuits of high-voltage d-c systems.[3] This new type of circuit can be used advantageously in the majority of the conventional metering schemes since the performance stability of the basic saturable reactor is combined with an effective to-zero operating characteristic. This new circuit also extends the use of this type of magnetic amplifier or transductor to include telemetering applications. The impedance matching of the transformer permits the load resistance to be of a value that can readily accommodate the transmission-line resistance and a dummy resistance which might be used to "swamp-out" transmission-line resistance changes.

Conclusion

The practical value of this circuit is its ability to operate with a to-zero characteristic similar to that of a conventional push-pull circuit while retaining the performance stability and simplicity of auxiliary circuitry which is associated with the basic single-ended circuit. The field of d-c metering is a natural area of application for a device having this combination of characteristics since such a device can act as an isolating full-range linear link between a high-voltage d-c power system and its associated metering and control circuits.

References

1. SERIES-CONNECTED SATURABLE REACTOR WITH CONTROL SOURCE OF COMPARATIVELY LOW IMPEDANCE, **H. F. Storm.** *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 756–65.

2. COMPENSATING FOR THE QUIESCENT CURRENT IN MULTISTAGE MAGNETIC AMPLIFIERS, **Alan S. FitzGerald.** *AIEE Transactions*, vol. 71, pt. I, Jan. 1952, pp. 7–12.

3. MAGNETIC-AMPLIFIER APPLICATIONS IN D-C CONVERSION STATIONS, **W. A. Derr, E. J. Cham.** *AIEE Transactions*, vol. 72, pt. III, April 1953, pp. 220–29.

Measurements of Materials with High Dielectric Constant and Conductivity at Ultrahigh Frequencies

H. P. SCHWAN
ASSOCIATE MEMBER AIEE

KAM LI
NONMEMBER AIEE

THE principles for measuring dielectric properties of materials at ultrahigh frequencies are well established.[1,2] They are based on measurement of voltage standing wave ratio (VSWR) and location of the standing wave pattern, which results from reflection of electromagnetic waves in front of the sample; see Fig. 1. Transmission line or wave guide sections, depending upon frequency, are used to obtain the necessary 1-dimensional field propagation. The complex reflection factor p, obtained from the measurement as indicated, is related to the dielectric properties of the sample. This relationship permits expressions of the dielectric properties directly, in terms of the observed quantities $W=1-\rho/1+\rho$ (inverse of VSWR) and $l_0=x_0+\lambda/4$. For the two most important special cases, sample loaded with either infinite or zero impedance, the relations given in Fig. 2 hold.[2] ϵ^* is the complex dielectric constant of the sample, ϵ its permittivity, and tan δ characterizes the dielectric losses. It is assumed that the dimensions of the transmission line are unchanged in the sample section, as indicated in Figs. 1 and 2.

This technique proved especially successful for materials with complex dielectric constants, whose magnitude is not too large compared with unity. However, no detailed discussion is given for the application of this technique to the measurement of materials with high complex dielectric constant. Interest in the measurement of such material has increased recently, partially because of the development of dielectrics with high dielectric constant and because of interest in the dielectric properties of biological material.[3-5] The purpose of this article is to summarize how measurements of dielectric properties of material with high complex dielectric constant are performed to obtain accurate results throughout the frequency range from 100 to 1,000 megacycles (mc).

Problem of Sample Thickness

It is obvious from the equations in Fig. 2 that the difference between the two measurable quantities L_1 and L_2 becomes larger as the dielectric constant ϵ^* increases. It is furthermore seen that for large values of ϵ^*, tan L_1 itself becomes large unless the sample thickness d is very

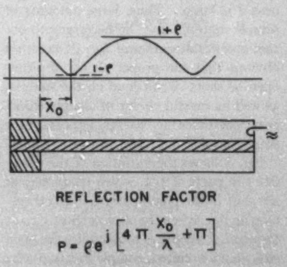


Fig. 1. Standing wave pattern in front of dielectric sample

Paper 54-524, recommended by the AIEE Electrical Techniques in Medicine and Biology Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted May 5, 1954; made available for printing September 8, 1954.

H. P. SCHWAN and KAM LI are with the University of Pennsylvania, Philadelphia, Pa.

This work was supported by the Office of Naval Research *Contract No. Nonr-551(05)* and by the Aeromedical Equipment Laboratory, United States Naval Base, Philadelphia, Pa.

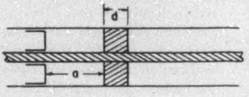


Fig. 2 (left). Sample loaded with infinite or zero impedance and equations for the determination of its dielectric properties

"OPEN" END $a = \frac{\lambda}{4}$:

$$\sqrt{\varepsilon^{+}}\ \tan\left[\frac{2\pi d}{\lambda}\sqrt{\varepsilon^{+}}\right] = \tan L_1 \quad (1)$$

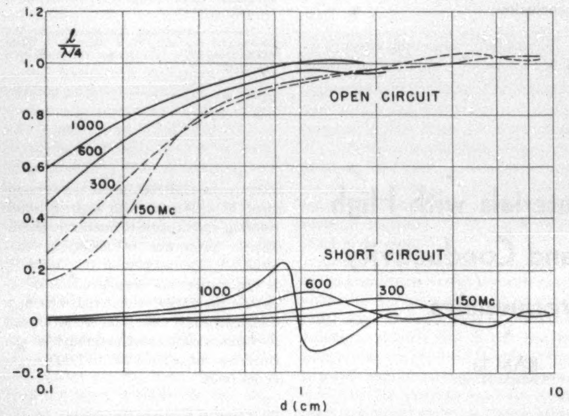
"SHORT CIRCUIT" END $a = \frac{\lambda}{2}$:

$$\frac{1}{\sqrt{\varepsilon^{+}}}\ \tan\left[\frac{2\pi d}{\lambda}\sqrt{\varepsilon^{+}}\right] = \tan L_2 \quad (2)$$

WHERE

$$L = \frac{2\pi \ell}{\lambda} - j \tanh^{-1}\frac{1-\rho}{1+\rho};\ \varepsilon^{+} = \varepsilon\,(1 - j\tan\delta)$$

Fig. 3 (below). Phase constant l as function of sample thickness and frequency for open- and short-circuited sample case; permittivity of sample 60, conductivity 0.01 ohm^{-1}cm^{-1}



small. Large values of tan L_1 establish undesirable operating conditions since a small error in the measured argument L_1 will yield large errors in ϵ^* because of the sensitiveness of the tan L_1 function to its argument. Tan L_2, on the other hand, is likely to be small unless the sample thickness d is large. Thus, large amounts of sample material may be necessary to obtain observable values of L_2. It becomes obvious that the proper choice of either open or short-circuit load on the sample, as well as careful choice of sample thickness d, becomes increasingly important for large complex dielectric constant values.

Fig. 3 shows the variation of the quantity l in parts of $\lambda/4$, which characterizes the shift in the phase pattern in the transmission line, when the open or short-circuit terminal of the line is replaced by the open and short-circuited sample respectively. It is plotted as a function of sample thickness for both open and short-circuited cases for a dielectric constant with a permittivity of 60 and an electric conductivity of 0.01 ohm^{-1}cm^{-1}. (These values are characteristic for tissues with high water content, body fluids, and water in the frequency range from 100 to 1,000 mc.) The maximum of l which can be obtained with the short-circuit technique over most of the d-range is considerably smaller than 1 cm. This means that an accuracy in l of 1 per cent requires exact readings to better than 0.1 millimeter (mm). This is difficult to achieve. Using the open-circuit technique the values of l approach $\lambda/4$ for all sample thickness greater than 1 cm. This will yield operating conditions with high sensitiveness of tan L_1 to its argument. Suitable values of l exist only for d-values of 1 to 2 mm. From this it follows that accurate determinations of ϵ are possible only with d-values which are rather small.

In Fig. 4 the inverse of VWRS is plotted against d. In the open-load case, W-values are obtained which vary around 0.1. They do not change rapidly with d as long as d is smaller than 1 cm. Rapid variations of W with d occur at higher frequencies for d-values above 1 cm. They are undesirable since they require extreme accuracy in the d-determination to avoid errors. In the short-circuit case W varies rapidly with d for smaller d-values, and approaches values of less than 0.001. Such W-values are also undesirable for measurement, since they compare with the W-values caused by the losses of the transmission line equipment itself (VSWR about 1,000). Complicated relationships exist for d greater than 1 cm, with more or less pronounced peaks, which in actual operation would again necessitate extreme accuracy in d. Therefore the W-values plotted in Fig. 4 substantiate the conclusion that most desirable operation is obtained with small d-values and unloaded sample.

If the sample thickness becomes sufficiently large, the input impedance of the sample becomes practically identical with the characteristic impedance of the sample section. In this case, it does not matter if the sample is loaded with either infinite or zero impedance. We will obtain W- and l-values which do not change when the initially unloaded transmission line is loaded by either the unloaded or short-circuited sample. We expect, therefore, that W_1 and W_2, as well as l_1 and $l_2+\lambda/4$, i.e., L_1 and $L_2+\pi/2$ become identical. This tendency is noticeable in Figs. 3 and 4. It is obvious at the same time that a 10-cm sample is only a rough approximation of an infinite d-load since W still varies about 20 per cent with d and load. Equations in Fig. 2 reduce for infinite d to

$$\sqrt{\epsilon^*} = j\tan L_1;\ 1/\sqrt{\epsilon^*} = j\tan L_2 \qquad (1)$$

which separate into

$$\epsilon = \frac{W^2\left[1+\tan^2\left(\frac{2\pi l}{\lambda}\right)\right]^2 - \tan^2\left(\frac{2\pi l}{\lambda}\right)[1-W^2]^2}{\left[1+W^2\tan^2\left(\frac{2\pi}{\lambda}\right)\right]^2} \sim \frac{1}{W^2} -$$

$$\epsilon\tan\delta = \frac{2W(1-W^2)\tan\left(\frac{2\pi l}{\lambda}\right)\left[1+\tan^2\left(\frac{2\pi l}{\lambda}\right)\right]}{\left[1+W^2\tan^2\left(\frac{2\pi l}{\lambda}\right)\right]^2} \sim \frac{2}{W^2\tan\left(\frac{2\pi l}{\lambda}\right)} \qquad (2)$$

This shows that it is possible to obtain good measurements of the permittivity from W. The losses characterized by $\epsilon \tan \delta$ are difficult to obtain for materials with high complex dielectric constant. This follows from the fact that they change with $\tan 2\pi l/\lambda$, which is considerably larger than 1, and are therefore sensitive to small changes of l.

Infinite sample thickness, short- and open-circuited sample, represent the only cases where relatively simple relationships are obtained. In all other cases of sample load the relationships are so complicated that they are not suitable for evaluation. The foregoing discussion therefore proves that small sample thickness and open-circuit technique provide the only feasible possibility for determining dielectric properties of material of high complex dielectric constant in the frequency range from 100 to 1,000 mc.

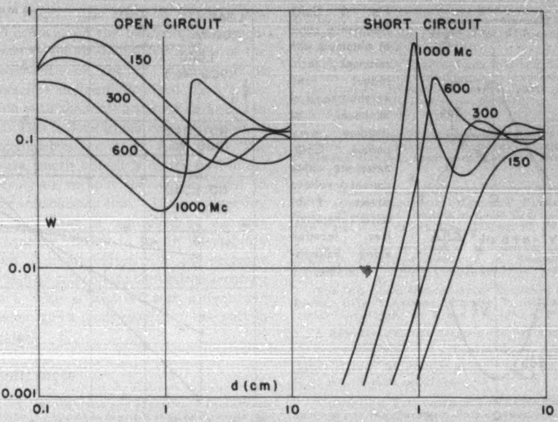


Fig. 4. Inverse of voltage standing wave ratio (W) as function of sample thickness and frequency for open- and short-circuited sample case; sample permittivity 60, conductivity 0.01 ohm^{-1}cm^{-1}

Method of Evaluation

Determinations of the complex dielectric constant based on either of the two equations in Fig. 2 necessitate the availability of complex tables of the functions $w = z \tan z$ and $w = 1/z \tan z$. The first of these two functions does not exist in tabulated form, while the second, applying to the undesired short-circuited sample case, has been given by Kennelly.[6] The desirability of small d-values fortunately makes it possible to use approximations which are obtained if $\tan (2\pi d/\lambda\sqrt{\epsilon^*})$ is replaced by its argument. This is permissible, especially in the case of large values of ϵ^*, as long as the magnitude of $\tan L_1$ is not very much greater than unity. The equations obtained are

$$\frac{2\pi d}{\lambda}\epsilon = \tan \frac{2\pi l}{\lambda} \frac{1-W^2}{1+W^2 \tan^2\left(\frac{2\pi l}{\lambda}\right)} \sim \frac{2\pi l}{\lambda}$$

$$\frac{\epsilon 2\pi d}{\lambda} \tan \delta = W \frac{1+\tan^2 \frac{2\pi l}{\lambda}}{1+W^2 \tan^2\left(\frac{2\pi l}{\lambda}\right)} \sim W \qquad (3)$$

The error of these equations may be estimated by use of the first correction term in the series development

$$\epsilon^* = \frac{\lambda}{2\pi d} \tan L_1 \left[1 - \frac{1}{3}\frac{2\pi d}{\lambda} \tan L_1 + \frac{4}{45}\left(\frac{2\pi d}{\lambda}\right)^2 \tan^2 L_1 - \ldots\right]$$

and is usually small. It is given by the equations

$$\frac{\Delta\epsilon}{\epsilon} = \frac{1}{3}\left(\frac{2\pi d}{\lambda}\right)^2 \epsilon (1 - \tan^2 \delta)$$

$$\frac{\Delta(\epsilon \tan \delta)}{\epsilon \tan \delta} = \frac{2}{3}\left(\frac{2\pi d}{\lambda}\right)^2 \epsilon$$

[$\Delta\epsilon/\epsilon$ and $\Delta(\epsilon \tan \delta)/\epsilon \tan \delta$ relative error of equations 3]

If l becomes so small that $\tan (2\pi l/\lambda)$ can be replaced by its argument and if $W^2 << 1$ the equations $\epsilon \sim l/d$ and $\epsilon \tan \sigma \sim \lambda/2\pi d\, W$ result.

Approximations for small d-values which hold in the case of short-circuit samples are

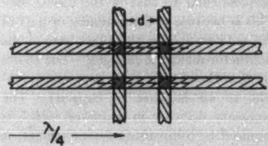
$$l = \frac{1}{3} d \left(\frac{2\pi d}{\lambda}\right)^2 \epsilon$$

$$\tan \mathrm{h}^{-1} W = \frac{1}{3}\left(\frac{2\pi d}{\lambda}\right)^2 \epsilon \tan \delta$$

They point out one of the major disadvantages of the short-circuit technique, i.e., both l and W are functions of the third power of d and either second or third power of the frequency. In the open-circuit case, on the other hand, both l and W are in a first approximation linearly dependent on sample thickness and frequency. This makes it possible to operate with one sample thickness in the open-circuit case over one total frequency decade, an advantage which is important when variation of d is either difficult to accomplish or undesirable.

Sample Holder Construction

The measurement of dielectric properties of solutions requires the use of a sample holder, which provides a layer of solution of thickness d and boundaries perpendicular to the conductors. Such a sample holder must be constructed so that its influence on the observed quantities l and W, and finally on the values ϵ^*, can can be understood easily. The best solution is the use of insulators which are sturdy enough to avoid bulging near the conductors, as thin sheets of plastic material do under pressure. Plates of polystyrene of some mm thickness are best suited. It seems logical to change



$$ZG = W \frac{1+\tan^2 2\pi \frac{l+l_0}{\lambda}}{1+W^2 \tan^2 2\pi \frac{l+l_0}{\lambda}}$$

$$Z\omega C = \tan 2\pi \frac{l+l_0}{\lambda} \frac{1-W^2}{1+W^2 \tan^2 2\pi \frac{l+l_0}{\lambda}} - \tan 2\pi \frac{l_0}{\lambda}$$

Fig. 5. Effect of sample holder plates on evaluation of dielectric properties. The effect of the back plate is eliminated by adjustment of a λ/4 section with reference to the front boundary of the back plate (arrow). The capacity of the front plate is corrected as indicated by the equations

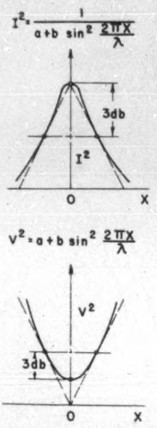

$I^2 = \frac{1}{a+b\sin^2\frac{2\pi X}{\lambda}}$
3db
I^2
O
X
$V^2 = a+b\sin^2\frac{2\pi X}{\lambda}$
V^2
3db
O
X

Fig. 6 (left). Resonance curve as obtained with resonance techniques and neighborhood of minimum of standing wave pattern. Ordinates are either current or voltage square. (Abscissa x characterizes location along transmission line)

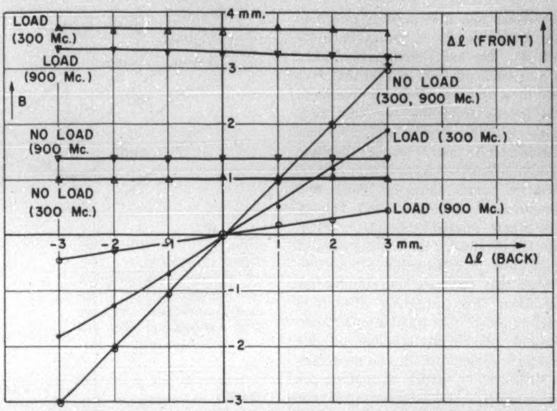


Fig. 7. Variation of phase constant l with adjustment λ/4 section in back of sample. Δl (front) is the change in phase constant l as caused by an error Δl (back) in the adjustment of the λ/4 section. Sample load: distilled water (d = 1.54 mm)

the diameter of the transmission line conductors while passing through the sample holder plates in such a manner that the characteristic impedance does not vary. This would avoid reflections at the surface of the sample holder plates and enable measurements without any consideration of the sample holder plates. However, a more detailed investigation shows that the reduction in conductor diameter, which is required to achieve equality of the characteristic impedances in the transmission line and plate, is considerable. It is larger than estimated, on the basis of a calculation which assumes that the orientation of the electric field is always perpendicular to that of the conductors. This is because of a shortening of part of the field lines near the edges, where the conductor diameter changes. The shortening of the electric field lines is equivalent to an increase in capacity. The necessary decrease in diameter, which is required to compensate for this effect, is often sufficiently great to reduce substantially the mechanical stability of the conductor arrangement.

Another possibility which we have chosen uses transmission line conductors passing through the sample holder plates, without change in diameter. Both back and front plate of sample act in this case as lumped capacitors as long as their thickness is very small compared to the electrical wave length. The influence of the back plate is adjusted for by taking into account that each mm of the plate replaces ϵ_p mm transmission line section (ϵ_p permittivity of sample plates) and by adjusting the λ/4 section in back of the sample, which is required for infinite load, to the front of the back plate, as indicated in Fig. 5. The capacity of the front plate has a value which may be considered in form of a quantity $l_0 = (\epsilon_p - 1)D$ where D is the thickness of the plate. It has been shown[7] that the effect of this quantity is corrected in any measurement of conductance G and capacity C when the quantity l in equations 3 is replaced by $l - l_0$ and when the product $Z\omega C$ or the value $2\pi d/\lambda\epsilon$ thus obtained is reduced by $\tan 2\pi l_0/\lambda$. (See Fig. 5.) It should be pointed out that neglect of this correction in l, even though it is small, may yield major errors in ϵ^* if not considered. This is especially true whenever ϵ^* is large enough so that $\tan 2\pi l/\lambda$ becomes more than linearly dependent upon its argument.

Accuracy of Measurement

The accuracy with which final results are obtained, depends upon the accuracy of electrical registration and mechanical positioning. The latter is, in general, about 0.1 mm and can be increased to 0.01 mm with micrometer drive. Registration of current and voltage amplitudes may be assumed accurate within about 1 per cent under favorable conditions. This is a much more severe limitation than those on mechanical grounds, at least in all such cases where the dielectric losses are large enough so that W exceeds a value of about 0.01. In Fig. 6, resonance peak, obtained by impedance measurements employing resonance techniques[8,9] and voltage minimum, as observed by measurement of the voltage pattern in front of the dielectric sample,[1] are demonstrated. In both cases the width of the resonance curve or minimum, as determined 3 decibels (db) below peak or above minimum and characterized by B, are identical with each other. They are related to W by the equation

$$\frac{1}{W^2} = 1 + \frac{1}{\sin^2\frac{\pi B}{\lambda}} \sim \left(\frac{\lambda}{\pi B}\right)^2$$

The accuracy of the value B is therefore, for practical purposes, Fig. 6(A), identical with the accuracy obtained in W. In the resonance curve method, Fig. 6(A), the tangent at the 3-db points passes through the resonance peak whenever $\sin^{-1} 2\pi x/\lambda$ can be replaced by $2\pi x/\lambda$. Under the same assumption the tangent through the 3-db points of the voltage pattern, Fig. 6(B), passes through the abscissa for $X = 0$. The relative accuracy in the bandwidth is therefore identical with the accuracy in the amplitude readings. As a result, the accuracy in W is about 1 per cent.

The accuracy in l, i.e. the accuracy with which peak of resonance curve or minimum of voltage pattern can be detected is derived as follows: Peak or minimum are most accurately determined by the 3-db points. This follows from the fact that the curves are symmetrical and that the change in amplitude with x is pronounced at the 3-db points, while it is zero in peak and minimum. The absolute accuracy in the peak and minimum location is therefore given by the accuracy in amplitude reading and the bandwith. The relative accuracy in l is identical with the relative accuracy in bandwidth multiplied by the ratio $2l/B$. If the dielectric constant and losses are such that the approximations in equations 3 can be used, the relative accuracy of permittivity ϵ is tan δ per cent, while the relative accuracy of ϵ tan δ is about 1 per cent.

Of further interest is the accuracy required in the adjustment of the $\lambda/4$ section in back of the sample holder. Fig. 7 shows that the effect of an erroneous setting of this $\lambda/4$ section appears more or less reduced in the corresponding change in front of the sample. This is understandable in view of the fact that only part of the energy, which hits the sample, passes through to the $\lambda/4$ section. Both error in $\lambda/4$ setting and change in l are, of course, identical if the sample holder is not loaded with material, as shown in the graph. The effect of errors in $\lambda/4$ adjustment on the bandwidth are negligible as shown in the dependence of the bandwidth on the $\lambda/4$ setting. It is therefore sufficient to adjust the $\lambda/4$ section with the same accuracy as desired in l. This is done in our case with a set of micrometers exact to 0.03 mm. The wave length λ itself is determined in experiments without the sample holder in the transmission line. The oscillator frequency must be controlled by accurate frequency meters to insure its stability and reproducibility with an accuracy which corresponds to the desired accuracy in l. This is achieved with a frequency meter which reads exactly to one part in 10,000.

References

1. A New Method for Measuring Dielectric Constant and Loss in the Range of Centimeter Waves, **S. Roberts, A. V. Hippel.** *Journal of Applied Physics*, New York, N.Y., vol. 17, 1946, pp. 610–16.

2. Methods of Evaluation of Dielectric and Magnetic Properties of Matter, **Herman P. Schwan.** *Annalen der Physik*, Leipzig, Germany, vol. 5, nos. 6–8, 1950, pp. 287–310.

3. Application of Electric and Acoustic Impedance Measuring Techniques to Problems in Diathermy, **Herman P. Schwan, Edwin L. Carstensen.** *AIEE Transactions*, vol. 72, pt. I, May 1953, pp. 106–10.

4. Heating of Fat-Muscle Layers by Electromagnetic and Ultrasonic Diathermy, **Herman P. Schwan, Edwin L. Carstensen, Kam Li.** *Ibid.*, Sept., pp. 483–88.

5. Capacity and Conductivity of Body Tissues at Ultrahigh Frequencies, **Herman P. Schwan, K. Li.** *Proceedings*, Institute of Radio Engineers, New York, N.Y., vol. 41, Dec. 1953, pp. 1735–40.

6. Chart Atlas of Complex Hyperbolic and Circular Functions (book), **A. E. Kennelly.** *Harvard University Press*, Cambridge, Mass., 1924.

7. The Influence of Supporting Elements at Lecher Line Terminals, **Herman P. Schwan.** *Annalen der Physik*, vol. 5, nos. 6–8, 1950, p. 268.

8. A Resonance Curve Method for the Absolute Measurement of Impedance at Frequencies of the Order 300 Mc/Sec, **R. A. Chipman.** *Journal of Applied Physics*, New York, N.Y., vol. 10, 1939, pp. 27–38.

9. Theoretical Treatment of Resonance Procedures for the Determination of Impedances at Decimeterwaves, **H. P. Schwan.** *Annalen der Physik*, Leipzig, Germany, vol. 5, nos. 6–8, 1950, pp. 253–267.

A Magnetically Regulated Portable Battery Charger

R. E. D. ANDERSON
ASSOCIATE MEMBER AIEE

IN CONNECTION with the development of military communications equipment for field application, a set of electrical and mechanical requirements was established for a regulated power supply for battery charging and floating service which was particularly unusual for equipment of this type. Mechanically, a portable unit was required with the size and weight to be kept to a minimum. At the same same, extreme ruggedness and reliability were necessary to meet the rough treatment anticipated in field operation.

Electrically, the charger was to operate from single-phase power services of 95 to 130 or 190 to 260 volts, 50 or 60 cycles. A nominal d-c output of 50 volts at 10 amperes was required for floating a 24-cell high-gravity storage battery, and an output of 60 volts at reduced current was required for applying an equalizing charge.

Extreme accuracy of voltage regulation could not be justified, since the life of a high-gravity battery is not related to the stability of the floating voltage to any great degree. However, reasonably good output voltage regulation was required for variations of ±5 per cent from the average in input voltage, ±3 per cent in input frequency, load variations from no load to full load, and ambient temperature variations from −10 to +50 degrees centigrade. The output noise level at the battery was required to be less than 20 decibels (above reference noise, adjusted for *F1A* line weighting) with three chargers operating in parallel at rated full load. The charger was required to be self-protecting against current overloads resulting from low battery voltages following a primary power failure.

The stress on ruggedness and reliability indicated that a design combining a selenium rectifier with magnetic control would be most suitable. A rectifier curcuit of this type had been developed previously for Bell system application in charging and floating batteries associated with *TD-2* microwave radio repeater stations.[1] Experience with the earlier design indicated that the specified requirements of load regulation, ambient temperature range, overload protection, and filtering could be met quite readily.

The only serious drawback to the earlier design lies in the ferroresonant type of line regulation employed. This circuit is inherently sensitive to changes in input frequency, with the output voltage varying from about $1^1/_2$ to 2 per cent for a 1-per-cent change in frequency. The portable charger would be supplied in most cases from engine alternators having relatively poor frequency regulation as compared to commercial power. The design of a new line regulator circuit less sensitive to input frequency changes was therefore undertaken.

The circuit used, as shown in Fig. 1, is an adaptation of the previous design, utilizing a self-saturating magnetic amplifier for line regulation in place of the ferroresonant line regulator, and incorporating several other features to permit operation under the wide range of input

---

Paper 54-416, recommended by the AIEE Metallic Rectifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted July 20, 1954; made available for printing August 20, 1954.

R. E. D. Anderson is with Bell Telephone Laboratories, Inc., New York, N. Y.

The principle of the line regulator portion of the circuit was originated by D. H. Smith. The mechanical design of the charger and the coil designs were the responsibility of R. R. Gay and A. B. Haines respectively and their associates.

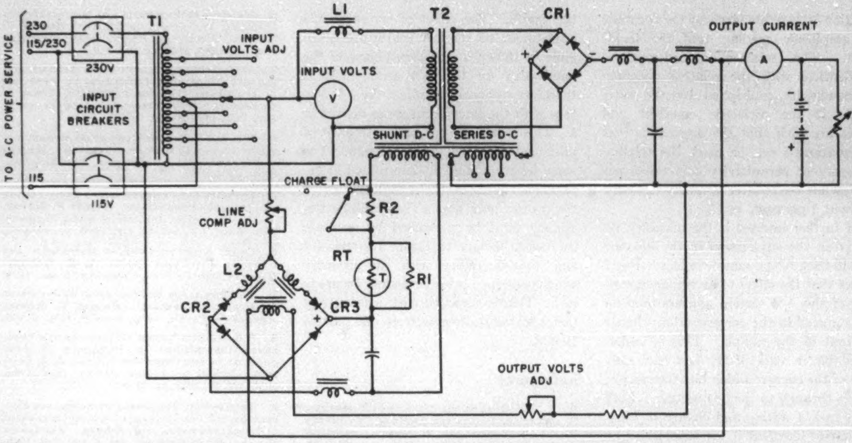


Fig. 1. Battery charger circuit

voltage and frequency and to provide the two output voltages required.

Description of Operation

To accommodate the extreme variations in line voltage, autotransformer *T1* permits operation from nominal inputs of either 115 or 230 volts merely by the closing of the proper input circuit breaker. Average input voltages anywhere in the ranges 95 to 130 or 190 to 260 volts are further adjusted to the optimum value of about 115 volts by means of the several taps on the autotransformer in conjunction with the tap switch and input voltmeter. The line regulating circuit, as described later, then need only compensate for fluctuations of ±5 per cent from the average value of line voltage.

The basic regulating circuit conists of a saturable transformer *T2* connected in series with a fixed ballast reactor *L1* across the adjusted a-c input line. The secondary windings of the saturable transformer supply power to the main rectifier and filter circuits which are thereby insulated from the input power services. Regulation is achieved by varying the a-c impedance of *T2* through control of the currents in its d-c windings. Neglecting wave form and phase relationships for simplicity, the input voltage applied to *L1* and *T2* in series may be considered as dividing across each in proportion to their impedances. The voltage applied to the main rectifier circuit may then be considered to be directly proportional to the impedance of *T2*.

The "shunt d-c" and "series d-c" windings on *T2* are oppositely poled. The charger load current passes through the series winding while the current in the shunt winding is derived from the line regulator circuit. With no external load connected to the charger, the d-c ampere-turns caused by the shunt winding predominate over the ampere-turns of the series winding by

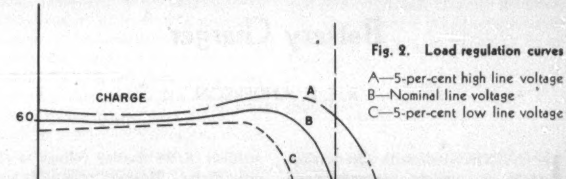


Fig. 2. Load regulation curves

a considerable amount. The net d-c flux saturates the core of *T2* and results in a relatively low impedance. As load current is drawn from the charger the series ampere-turns increase and, since they oppose the shunt ampere-turns, the net d-c flux decreases. The resulting increase in the a-c impedance of *T2* with load provides increased secondary voltage to the rectifier stack to compensate for the increased voltage drops in the rectifier and filter. Taps on the series winding are provided to adjust the degree of load compounding to accommodate manufacturing variations and aging of the selenium stack *CR1*. To simplify parallel operation of chargers, the taps are adjusted so that the unit is somewhat undercompounded.

With nominal input line voltage and at a load current slightly in excess of the rated full load, the series and shunt ampere-turns become equal and the net d-c flux is zero. As the load current is increased beyond this point, the circuit acts to reduce the output voltage sharply to protect the charger from harmful overload. Since the series ampere-turns now predominate, an increase in load current causes an increase in net d-c flux. The impedance of *T2*, therefore, decreases and the resulting loss in voltage applied to the rectifier, together with the increased filter losses, causes the output voltage to be reduced rapidly. In addition, the reduction of output voltage acts on the line regulator circuit to reduce the shunt ampere-turns, thereby increasing the net d-c saturation even further and resulting in practically constant current operation, as can be seen in Fig. 2. This feature permits an unattended charger to be connected to a discharged battery after a period of a-c power failure without damage to the charger from overloading.

Line Regulation

The line regulator circuit consists essentially of the self-saturating magnetic amplifier circuit, made up of regulator *L2* and rectifier stacks *CR2* and *CR3*. The output is filtered and passes through the float-charge switch and a temperature-compensating network to supply d-c current for the shunt winding of saturable transformer *T2*.

The current in the d-c control winding on *L2* is fed back from the charger output through potentiometer designated "output volts adj." The control winding is poled so that its ampere-turns aid the ampere-turns caused by self-saturation in the load windings.

The line regulator circuit is connected across the adjusted a-c input voltage to the main rectifier. Line voltage swings, therefore, result in variations in the ampere-turns in the shunt winding of *T2*. As the line voltage increases, the voltage across both the ballast reactor *L1* and the primary winding of *T2* would normally increase. However, the increased line voltage increases the net d-c flux in *T2* as a result of the increase in shunt ampere-turns. The corresponding reduction in a-c impedance of *T2* causes most of the increase in line voltage to be absorbed across the ballast reactor and the voltage applied to the main rectifier remains substantially constant. When the line voltage decreases, the opposite effect occurs, with the drop across the ballast reactor being reduced to keep the transformer voltage constant.

The degree of line compounding is adjusted by varying the potentiometer designated "line comp adj." For a given change in line voltage the change in shunt ampere-turns in *T2* is reduced as the resistance of the potentiometer is increased.

As noted in the foregoing, the net d-c flux in *T2* decreases with load current, because of the increase in current in the series winding. Since the changes in shunt current are substantially independent of the load, the effect of the line regulator on the net flux becomes greater as the load increases. The "line comp. adj." potentiometer is, therefore, adjusted to give optimum line compounding at about half load, with the circuit being somewhat undercompounded for line changes at lesser loads and overcompounded as the load is increased.

The signal fed back from the charger output to the control winding of *L2* acts to aid regulation of the output voltage for both line and load changes, e.g., an increase in output voltage causes an increase in d-c ampere-turns on *L2* due to the control winding. Since the control winding is poled to aid the self-saturation of the load windings, this has the effect of reducing the impedance of the load windings and of increasing the output current of the line regulator circuit. At charger load currents of less than rated full load, the resulting increased current in the shunt winding of *T2* produces a greater net d-c saturation of its core and corresponding reduction in the charger output voltage. Similarly, the reduced current in the control winding of *L2* resulting from a drop in output voltage tends to raise the output as long as the d-c ampere-turns in the shunt winding of *T2* predominate over the series winding. At loads in excess of rated full load, the decreased shunt winding current increases the net d-c saturation of *T2* and increases the

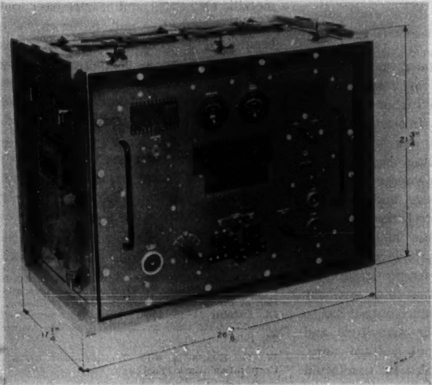
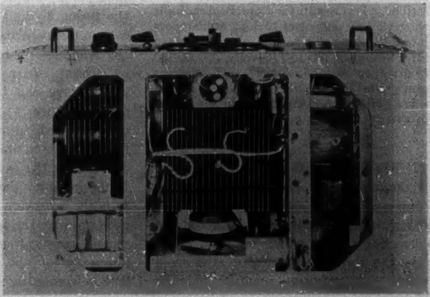


Fig. 3 (left). Front view

Fig. 4 (below). Top view with carrying case removed



slope of the "droop" portion of the regulation curve.

Variations in ambient temperature tend to cause an inverse variation in the shunt current in *T2* caused by changes in the resistance of the windings of *L2* and *T2*. The charger output voltage therefore tends to rise with an increase in ambient temperature, caused by the decreased shunt current and resulting increased a-c impedance of *T2*. To offset this effect, the negative temperature coefficient of thermistor *RT* is utilized in the shunt current loop. The parallel resistor *R1* adapts the thermistor characteristics to the needs of the circuit.

Operation of the float-charge switch to the charge position removes a short circuit from the *R2* resistor which is also in the shunt current loop. The resulting decrease in shunt current increases the a-c impedance of *T2* and raises the charger output voltage to a nominal 60 volts for use in applying an equalizing charge to the battery. Finer adjustment of the output voltage in both the float and charge conditions is supplied by the "output volts adj" potentiometer in series with the d-c control winding of *L2*.

Mechanical Design

A welded aluminum chassis is utilized to provide the ruggedness required with minimum weight. To keep the rectifier stacks as small as possible, forced air cooling is employed. Also, all coils are designed with class-*H* insulation and temperature rises run as high as possible, consistent with the circuit requirements, to minimize the weight of the charger. The chassis is fastened to a field-carrying case through a shock mount framework, and the complete assembly has successfully been subjected to severe shock and bounce testing. The combined weight of the charger and carrying case is 265 pounds. The overall dimensions are $21^{3}/_{4}$ by $26^{7}/_{8}$ by $17^{1}/_{4}$ inches. Fig. 3 is a front view of the unit mounted in the carrying case and Fig. 4 is a top view showing the internal chassis construction.

Performance

As illustrated in Fig. 2, the output voltage regulation of the charger for float operation is less than ±4 per cent for loads from no load to 10 amperes, and input line voltage variations of ±5 per cent from any nominal values in the ranges from 95 to 130 or 190 to 260 volts, at either 50 or 60 cycles. In the charge condition, the reduced current in the shunt d-c winding of *T2* effectively results in an increase in the load compounding and a decrease in the line compounding, but the over-all regulation is within the same limit. The reduced current also causes the overload droop point to occur at a lesser load current than in float operation since this occurs when the shunt and series ampere-turns are equal. Line voltage variations have a similar effect in both the float and charge conditions as a result of corresponding changes in the shunt current. The relative steepness of the constant-current portion of the curves as compared to earlier designs nullifies the effect of the shift in the droop point.

A 2-position tap switch is provided for adjusting the circuit for operation from either 50- or 60-cycle supplies. The switch selects taps on the windings of *L1* and *L2*. For variations of ±3 per cent from either nominal frequency, the output voltage at rated loads varies less than ±3 per cent. This represents about a 2-to-1 improvement over the ferroresonant type of line regulator circuit.

The temperature compensation circuit limits variations in output voltage to less than ±1 per cent for ambient temperature variations of ±30 degrees from the optimum ambient of +20 degrees centigrade.

The components and techniques used in this design result in a charger which meets the electrical requirements of good output voltage regulation for extreme input conditions, and also the mechanical requirements of minimum size and weight consistent with dependability and ruggedness.

Reference

1. Automatic Regulation of Metallic Rectifiers by Magnetic Control, **D. H. Smith.** *AIEE Transactions*, vol. 71, pt. I, Jan. 1952, pp. 111–14.

Saturating Transformer Reference Circuit

W. G. EVANS
ASSOCIATE MEMBER AIEE

Synopsis: The theory of operation of a static magnetic reference device is described. Test results for a practical example show that an output of 15±0.2 milliampere into a 50-ohm resistive load can be obtained for an input of 100 to 140 volts, 360 to 440 cycles per second, from −55 degrees centigrade (C) to +100 C ambient temperature. The power input to the device is 5 watts.

IN REGULATING and control systems the quantity to be controlled is often compared to a reference quantity. The difference or error is used to actuate a correction system which tends to maintain the controlled quantity at a fixed value. The reference quantity may be some physical constant such as the modulus of elasticity of a spring, e.g., the spring holding the carbon stack in a carbon-pile regulator. In electric systems the reference quantity can be a voltage or current. In a voltage regulator the line voltage can be compared to a reference voltage or current and the error can be used to control the generator field in order to maintain the line voltage constant as the generator load is varied. If the input to the reference is supplied from the generator output or an auxiliary power source, the reference output has to remain constant as the magnitude and frequency of the input voltage vary. The reference must also be unaffected by the ambient temperature. Any variation in the reference output will cause the generator voltage to vary.

Gas discharge tubes, nonlinear heated metallic elements, nonlinear resistor and inductor elements, and batteries have been used as reference devices. This paper describes a magnetic reference device which provides relatively constant output current when the magnitude and frequency of the input voltage are varied. The reference is made from static components and requires no electron tubes or heated nonlinear metallic elements. Output of the reference remains relatively constant as the ambient temperature is varied.

Paper 54-388, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 17, 1954; made available for printing August 24, 1954.

W. G. Evans is with Westinghouse Electric Corporation, East Pittsburgh, Pa.

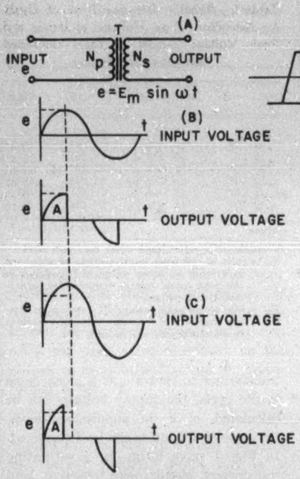


Fig. 1. Ideal saturating transformer wave forms—constant input frequency

Nomenclature

e = instantaneous value of wave form
E_m = peak value of input voltage
E_m' = peak value of input voltage referred to secondary
E_{m1} = peak value of wave form before saturation occurs
E_{m2} = peak value of input voltage wave form
E_{rms} = rms value of voltage wave form
E_p = peak voltage out of transformer for turns-ratio of 1
e_s = induced voltage in secondary winding
K_1 and K_2 = constants
N_s = number of turns on secondary
P = power, watts
R = resistance
R_L = resistance of load
R' = resistance in input of T referred to secondary
t = time, seconds
T = period of wave form (equation 4)
T = temperature, C (equation 15)
θ = angle for saturation
θ_1 = angle of cycle when saturation occurs
θ_2 = angle of half-cycle π radians
$\omega = 2\pi f$ (f: frequency, cycles per second)
ϕ = lines of flux linking N_s turns
ϕ_s = saturation flux of core

Saturating Transformer Characteristics

A basic part of the reference described here is a saturating transformer. In the ideal saturating transformer of Fig. 1(A) the voltage induced in the secondary is

$$e_s = N_s \frac{d\phi}{dt} 10^{-8} \qquad (1)$$

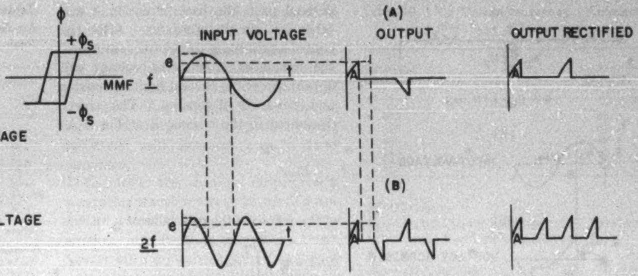


Fig. 2. Ideal saturating transformer wave forms—constant input voltage

Although the input voltage shown in Fig. 1(A) is sinusoidal, equation 1 holds for any input wave shape. Integration of equation 1 yields

$$\int e_s dt = N_s \Delta\phi 10^{-8} \qquad (2)$$

If the input voltage e is sufficient to change the flux in the core from $-\phi_s$ to $+\phi_s$ in Fig. 1(A), the area represented by $\int e_s dt$ will be constant even if the magnitude of the input e is varied. $\int e_s dt$ is constant because of the hysteresis loop assumed in Fig. 1(A). After saturation no flux change can take place in the core, and the right-hand portion of equation 2 becomes $N_s 2\phi_s 10^{-8}$ which is constant. It should be noted that even if the slope of the hysteresis loop of Fig. 1 is zero after saturation, air coupling exists between the input and output windings, but this air coupling has a negligible effect in the reference devices described here. The effect of varying the magnitude of the input voltage is shown in Fig. 1(B) and (C). As the input voltage increases, the core saturates sooner in the cycle of the input voltage, and the peak value of the output voltage increases. In Fig. 1(B) and (C) the areas A are equal.

It is assumed in the ideal case illustrated in Fig. 1 that no flux change occurs in the core after saturation and that exciting current is negligible. After saturation the input current would rise to an infinite value. In a practical transformer, resistance is placed in the input to limit current flow after saturation. The exciting current is negligible if the voltage drop, which occurs across this limiting resistor before saturation, is negligible.

When the frequency of the input voltage varies, equation 2 still holds if the input voltage is sufficient to saturate the core. The wave forms of Fig. 2 illustrate the effect of variable input frequency for the ideal case of Fig. 1. As frequency increases, saturation occurs sooner in time, and the peak value of the output increases when saturation occurs before $\pi/2$ radians. In each example the areas A are equal. When the output is rectified, there are more voltage pulses for a given time interval, and the average output will increase linearly with an increase in frequency. This action would be suitable for a frequency meter since the average output would vary linearly with frequency, and the output would be unaffected by variations in the magnitude of the supply. For a saturating transformer to provide constant average output for variable input frequency, the output must be a measure of area A or a frequency compensation system is necessary.

Although the output voltage of the saturating transformer is the quantity which would be used as a reference, the constancy of the output voltage is dependent upon the constancy of ϕ_s, saturation flux of a particular core. ϕ_s of ferromagnetic materials such as Hipernik V, Deltamax, and Orthonik will vary with temperature; this variation will be discussed later. Information

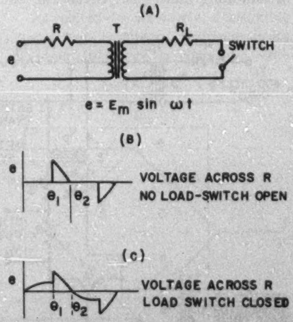


Fig. 3. Voltage wave forms across current limiting resistor

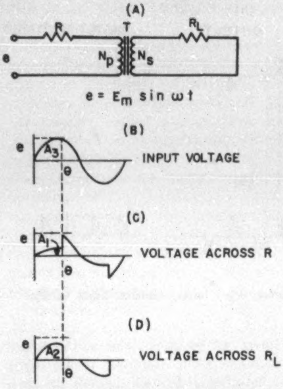


Fig. 4. Wave forms across current limiting resistor and load resistor

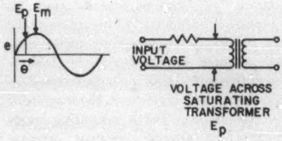


Fig. 5. Circuit for data of Table I

in the literature and experience indicates that for any given temperature in the range −60 C to +100 C ϕ_s for these materials will not vary with time in a manner that is observable in reference devices designed for 0.1-per-cent accuracy.

Saturating Transformer Design Factors

Equations necessary for the design of a saturating transformer circuit can be derived from the basic equation 1 and other a-c circuit principles. After the core of a saturating transformer saturates, the remainder of the line voltage will appear across the current limiting resistor and power is dissipated. The power dissipated in the resistor R of Fig. 3(A) is

$$P=\frac{E_{rms}^2}{R} \tag{3}$$

By definition the rms value of a wave is

$$E_{rms}=\sqrt{\frac{1}{T}\int_0^T e^2dt} \tag{4}$$

The typical wave form of voltage across the input resistance of an unloaded saturating transformer is shown in Fig. 3(B). Here again exciting current is considered negligible. The rms value of such a wave is

$$E_{rms}=\sqrt{\frac{1}{\theta_2}\int_{\theta_1}^{\theta_2} E_m^2 \sin^2 \omega\theta d\theta} \tag{5}$$

For example, if the transformer saturates at 90 degrees, then E_{rms} is $E_m/2$. As the transformer is loaded, voltage is dropped across R before the transformer saturates, and for a resistive load on the transformer a typical wave shape across R is as shown in Fig. 3(C). The rms value of such a wave form is

$$\sqrt{\frac{1}{\theta_2}\left(\int_0^{\theta_2} E_{m1}^2 \sin^2 \omega\theta d\theta + \int_{\theta_1}^{\theta_2} E_{m2}^2 \sin^2 \omega\theta d\theta\right)} \tag{6}$$

In the design of a saturating transformer, the supply voltage must be great enough to allow for the drop across R before the core saturates. Before saturation the transformer is a linear transformer, and current flows in the primary and secondary due to the load in the output. If it is desired to cause the transformer to saturate at a given angle in the cycle, the supply voltage can be calculated, of if the supply is known, the angle of saturation can be determined. In Fig. 4 wave forms for a saturating transformer circuit with resistive load are shown. From equation 1 and the transformer design, the volt-second area in the output is equal to $2N_s\phi_s10^{-8}$. In Fig. 4, areas A_1, A_2, and A_3 are volt-second areas appearing across the current limiting resistor, the load resistor, and the input terminals respectively

Table I. Relation Between Time of Cycle for Saturation, Peak Voltage of Input, and Peak Voltage Across an Ideal Unloaded Saturating Transformer

θ	E_m'	E_p'
$\pi/8$	13.2	5.1
$\pi/4$	3.4	2.4
$3\pi/8$	1.6	1.5
$\pi/2$	1.0	1.0
$5\pi/8$	0.72	0.67
$3\pi/4$	0.58	0.42
$7\pi/8$	0.52	0.20

where

θ = angle of input voltage when core saturates

Em' = ratio of the peak input voltage required for saturation at some value of θ radians to the peak input voltage required for saturation at $\pi/2$ radians

E_p' = ratio of the peak voltage E_p across a saturating transformer for some value of θ radians to E_p at θ equals $\pi/2$ radians

$$A_3-A_1=A_2 \doteq 2N_s\phi_s10^{-8} \tag{7}$$

(All quantities referred to secondary)

$$\int_0^{\theta}\frac{1}{\omega} E_m' \sin \omega\theta d\theta - \int_0^{\theta}\frac{1}{\omega}\frac{R'E_m'}{R'+R_L}\times \sin \omega\theta d\theta = 2N_s\phi_s10^{-8} \tag{8}$$

$$\frac{E_m'}{\omega}\left(\frac{R'}{R'+R_L}\cos\omega\theta - \cos\omega\theta - \frac{R'}{R'+R_L}+1\right) = 2N_s\phi_s10^{-8} \tag{9}$$

If R_L is a load made up of resistance, capacitance, and inductance, the waves shape across R will be different, but

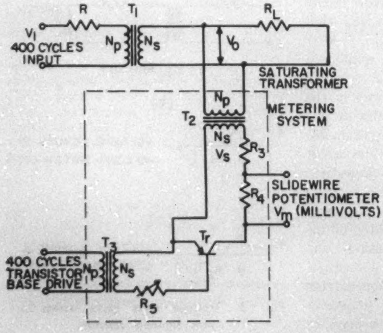


Fig. 6 (left). Average measuring metering system

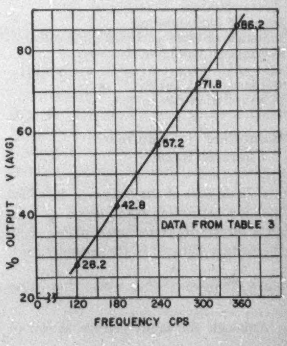


Fig. 7 (right). Output versus frequency—input voltage constant (57V$_{rms}$)

Table II. V_o Versus V_I for Various Values of R_L. Constant Frequency 400 Cycles per Second

V_I RMS	V_o Average ($R_L=\infty$)	V_o Average ($R_L=20{,}000$)	V_o Average ($R_L=2{,}000$)
50	93.50	88.50	54.34
60	102.80*	103.26*	71.10
70	103.34	103.34	83.02
90	103.42	103.36	103.14*
110	103.52	103.38	103.20
130	103.62	103.39	103.21
150	103.78	103.41	103.22

* Saturation occurs.

equation 7 can be used to find the input voltage or the angle of saturation. If the angle of saturation and E_m are known for an input voltage as in Fig. 4(B), the peak value of voltage on the output of an unloaded saturating transformer is simply

$$E_p = E_m \sin \omega\theta \tag{10}$$

Equation 10 is valid for θ equal to values up to $\pi/2$ radians. From $\pi/2$ to π radians, E_p is equal to E_m. When the transformer is loaded, E_p is reduced by the voltage drop across R, the resistance in the input of the saturating transformer.

Table I is based upon the circuit shown in Fig. 5. This table shows the relation between time of cycle for saturation, peak value of input voltage, and peak voltage across a saturating transformer calculated from equations 9 and 10. The table is for an ideal saturating transformer, but experimental data closely verify the table. Table I shows that the peak input voltage E_m must rise to a very high value as compared to E_m for θ equal to $\pi/2$ if the core is to saturate early in the input cycle. Since the power lost in the limiting resistor varies approximately as the square of the input voltage divided by the limiting resistance, the power lost in the limiting resistance will increase greatly as θ decreases.

To check the average output of a saturating transformer, an accurate measuring system is required. Since the average output of a saturating transformer will vary only slightly if frequency is constant, an expanded scale meter or a meter capable of indicating a change of 1 part in 1,000 is needed. Fig. 6 shows a system which may be used to measure the average value of a symmetrical chopped sine wave. The components of Fig. 6 are as follows:

T_1—saturating transformer.
R—current limiting resistor.
R_L—load.
T_2—impedance matching transformer (linear transformer); V_s to be less than 30 volts peak.
R_3—10,000 ohms.
R_4—100 ohms.
T_r—*CK722* Raytheon transistor.
R_5—750 ohms (base current limiting).
T_3—base drive transformer; V_3 to be 6 V_{rms}. Base voltage and input voltage to be synchronized.

A slide-wire potentiometer is used to measure the rectified half-wave output of the saturating transformer. T_r is a transistor operated as a synchronous switch to provide half-wave rectification. The transistor has a threshold voltage of about 1 millivolt and a forward resistance of about 10 ohms. Thus, the transistor provides accurate half-wave rectification, and the slide-wire potentiometer provides an accurate measure of the average output voltage. Tables II and III and Fig. 7 are compiled from data taken with the metering circuit shown in Fig. 6. The saturating transformer and current limiting resistor R connected as in Fig. 6 are as follows:

Table III. V_o Versus Frequency. Constant Voltage $57V_{rms}$

f, Cycles per Second	V_o Average ($R_L=\infty$)
360	86.2
300	71.8
240	57.2
180	42.8
120	28.2

R—200 ohms.
T_1—$1^1/_4$ by $1^3/_4$ by 1/2 Hipernik V toroidal core, 2-mil strip thickness.
N_p—300 *t*, no. 22.
N_s—750 *t*, no. 28.

V_o is the average value of the output voltage of the saturating transformer for various load resistances R_L and for N_p/N_s equal to 10 for T_2. From the metering circuit constants, the average value of V_o for the saturating transformer just described, neglecting losses in T_2 and T_r, is

$$V_o = \frac{V_M}{R_4}(R_3+R_4)\frac{N_p}{N_s}\frac{2}{1{,}000} = \text{average voltage} \tag{11}$$

$$V_o = V_M \times 2.02$$

where

$\frac{2}{1{,}000}$ = factor to convert V_M from millivolts to volts and half-wave readings to full-wave values

Table II gives data for V_o versus V_I, and Table III gives data for V_o versus frequency. Table II shows that after the core saturates the input voltage may increase 50 per cent while the average output will vary about 0.1 per cent. Table III and Fig. 7 show that the average output varies linearly with respect to the input frequency.

Frequency Compensation

As pointed out previously, a saturating transformer is a useful reference device if input frequency is constant. Since

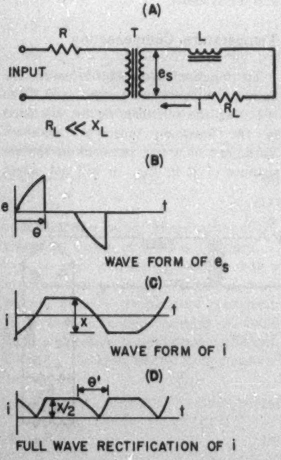


Fig. 8 (left). Inductance in output of a saturating transformer

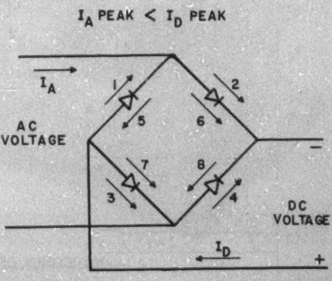


Fig. 9 (right). Biased rectifier

the average output will vary linearly with frequency, a linear inductor may be considered as a frequency compensating means as in a system in which the voltages and currents are sinusoidal. The current change in a linear inductor when resistance is neglected is

$$i = \frac{1}{L}\int e\,dt \qquad (12)$$

In Fig. 8 a saturating transformer, linear inductor, and associated wave forms are shown. It is easily shown mathematically or graphically that, since the value of $\int e\,dt$ is constant in a saturating transformer (see equation 2) for variable input voltage or frequency, the value x of Fig. 8(C) will be constant. As frequency or voltage vary, the angle of saturation θ in Fig. 8(B) will vary. In Fig. 8(D) the value θ' will vary, and consequently the average value will vary, even though x and $x/2$ are constant. If the average value of the rectified output current in Fig. 8 is to remain constant, the current during the interval θ' must remain constant.

Before considering a means to keep i constant during the interval θ', consider the performance of an ideal rectifier as shown in Fig. 9. If the rectifier is subjected to direct and alternating voltages, the alternating current I_A can pass through the rectifier as shown without impedance from the direct current I_D if the peak value of I_A is less than the peak value of I_D. It appears as if I_A passes through some of the rectifier cells in the blocking direction. However, the net rectifier current still flows through the rectifier in the forward direction. The rectifier appears as a short circuit to the direct and alternating currents. Rectifiers are commonly operated in this mode in regulator systems; rectifiers operated in this mode are termed biased rectifiers.[1] The linear inductance L_2 in the output circuit of Fig. 10 offers a means of maintaining i_2 relatively constant. On the curve of i_2, the dotted lines show the corresponding rectified values of i_1. During the interval of time 3 to 4, e is zero, and the total current in the circuit will decrease at a rate determined by the time constant of the system $(L_1+L_2)/R$, where R is the resistance of the inductors, the transformer secondary, the rectifier forward resistance, and the load. At time 4, e begins to change, and if the resistance of the system is again neglected, i_1 will change according to

$$i_1 = \frac{1}{L_1}\int e\,dt \qquad (13)$$

This happens because as e causes i_1 to decrease the peak value of i_1 is less than the peak value of i_2. The polarity of L_2 reverses from that shown in Fig. 10 in attempting to maintain i_2 constant, and the system meets the requirements for a biased rectifier. As long as the rectifier is biased, the peak value of i_1 is less than the peak value of i_2; the alternating system is isolated from the direct system. From 4 to 5, i_2 will decrease at a rate determined by the time constant L_2/R', where R' is the resistance of L_2, the load, and the rectifier forward resistance. At time 5, when i_2 and i_1 are again equal, the currents will change according to

$$i_1 = i_2 = \frac{1}{L_1+L_2}\int e\,dt \qquad (14)$$

At time 6, the voltage e ceases to change, and the cycle will repeat as before. If the time constant of the a-c side of the rectifier is shorter than the time constant of the d-c side, at some time just after 3 the peak value of the alternating current would be less than the peak value of the direct current, and the rectifier would be biased at some time between 3 and 4. The current in the d-c side of the rectifier would be determined by the time constant L_2/R'. An important consideration may be noted from the foregoing discussion. The time constants $(L_1+L_2)/R$ and L_2/R' should be as long as possible to improve the accuracy of the reference when voltage and frequency change. Since the interval 4 to 6 will vary as the frequency or magnitude of the input voltage varies, the average value of i_2 will vary unless the time constants $(L_1+L_2)/R$ and L_2/R' are long with respect to the period of the input voltage. The interval 4 to 6 may be shortened by causing the transformer to saturate early in the input cycle.

There may be some question about the use of resistance in the analysis of the circuit at some times and the neglect of resistance at other times. Resistance could be included in equations 13 and 14, but if the time constants of the a-c and d-c systems are long, the inductances largely determine the currents in the a-c and d-c systems.

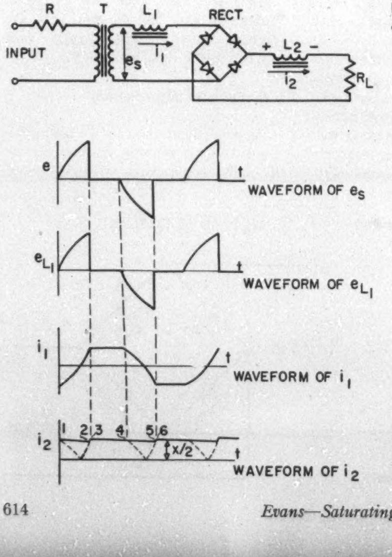


Fig. 10 (left). Inductance in rectified output

Temperature Compensation

To function satisfactorily over the ambient temperature range, −60 C to +100 C, the reference device discussed in the foregoing must be modified. Variations of a few per cent in the resistance of R in Fig. 10 will not affect

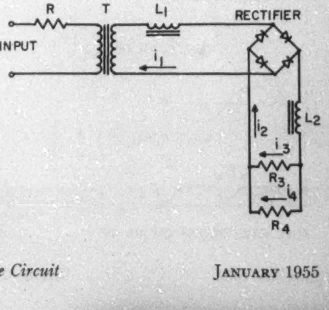


Fig. 11 (right). Temperature compensated reference

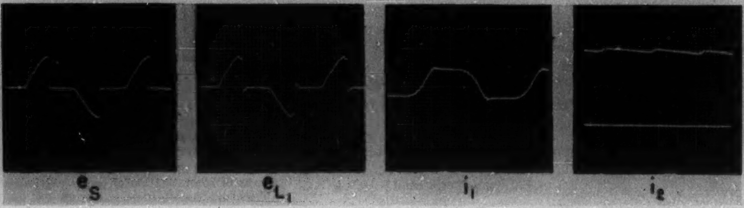


Fig. 12. Oscillographs of current and voltage wave forms for circuit of Fig. 10

the reference output. The forward and back resistance of commercially available rectifiers will change with temperature, but these changes have been found to have negligible effect upon the reference. Since the time constants of the a-c and d-c circuits are long, the small changes in the forward drop of the rectifiers make no difference in the output. The back leakage current is in the order of microamperes from −60 C to +100 C, and the change in leakage which does occur has negligible effect. The permeability of the inductors L_1 and L_2 can be stabilized to ±0.1 per cent. Copper resistance of the circuit has little effect on the reference output because the time constants are long. T causes the greatest change in the reference output. As noted previously in the discussion of saturating transformer design factors, the average output is directly related to the flux ϕ_s required for saturation of the core. If ϕ_s varies, the average output will vary. It has been found that the saturation flux of Hipernik V and Deltamax core material varies as the temperature of the core varies in a manner expressed by

$$\phi_s = K_1 - K_2T \tag{15}$$

ϕ_s will decrease about 0.07 per cent per C increase in temperature.

Fig. 11 shows a circuit which can be used to temperature compensate the reference output. i_2 decreases in the same manner as ϕ_s

$$i_2 = K_3 - K_4T \tag{16}$$

$$i_2 = i_3 + i_4 \tag{17}$$

If R_3 varies as the temperature of the reference varies and R_4 is constant, i_4 can be kept constant. Since i_2 has a negative temperature coefficient, R_3 must have a positive temperature coefficient. If R_3 were to remain constant, R_4 could have a negative temperature coefficient. Consider the case where R_4 is constant. In Fig. 11

$$i_3R_3 = i_4R_4 \tag{18}$$

$$R_3 = \frac{i_4R_4}{K_3 - K_4T - i_4} \tag{19}$$

Since the slopes of i_2 and ϕ_s are equal, 0.07 per cent per C, K_3 and K_4 are respectively 0.01965 and 14×10^{-6}. Equation 19 is not linear, but R_3 would have a positive temperature coefficient. The slope of R_3 increases as T increases in the range from room temperature to 100 C. Commercially available positive and negative temperature coefficient resistors are suitable for R_3 and R_4.

Reference Design Method

The design of a practical reference may be approached by the following steps:

1. Determine input magnitudes and variations. Determine the output current and load resistance.

2. Calculate L_2 in Fig. 10 to give L_2/R' about 50 times as great as the period of the input voltage. L_1 will have about the same value as L_2 for $(L_1+L_2)/R$ about 50 times the period of the input.

3. For ±10-per-cent voltage and frequency variations of the supply, assume saturation of the transformer T at $\pi/2$ radians for the first calculation. Determine from equation 13 the peak value of the transformer secondary voltage.

4. Design the transformer according to conventional practice. Design for saturation at half the nominal supply voltage and frequency. In cases where saturation is to occur at values different from $\pi/2$ radians, the transformer is designed to support a voltage in accordance with equation 2. Match the input voltage and output voltage by selection of the proper turns-ratio for T.

5. Determine the value of R for allowable power dissipation from equation 6. If R is large enough so that voltage dropped before saturation is important, check the design per equations 7, 8, and 9 to see that saturation of T is maintained at the lowest value of voltage and the highest value of frequency of the input.

6. Temperature compensate when necessary.

7. Check accuracy of design with variations in input voltage and frequency and variations in the ambient temperature.

It will be noted in the foregoing sections and in the design method just given that no equations are listed to determine the accuracy of the reference for a given size, weight, and input variation. The accuracy of the output is mostly a function of the time constants associated with L_1 and L_2 and the saturation angle

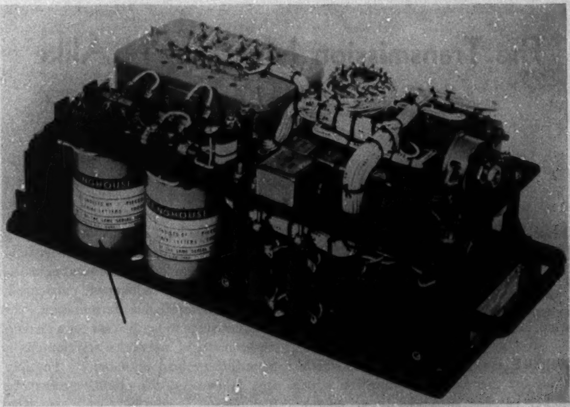


Fig. 13. Magnetic amplifier voltage regulator for a-c generator

of T for given characteristics of L_1, L_2, and T. Accuracy of a reference could be calculated as a function of these variables, but it is usually simpler to design the reference to give the longest possible time constants for a given size and weight. The earliest saturation of T for the power loss permitted will give maximum accuracy.

Practical Examples

A reference has been designed to illustrate the wave forms and accuracy obtained in a practical device. The characteristics of this device are as follows:

Input

360 to 440 cycles per second.
100 to 140 volts.
5 watts.

Output

Including variations when the ambient temperature varies from −55 C to 100 C. 15±0.2 milliampere d-c into a 50-ohm load.

The device has a volume of 15 cubic inches and weighs 1.5 pounds. Wave forms for the reference for the circuit of Fig. 10 are shown in Fig. 12. These wave forms closely resemble the theoretical wave forms of Fig. 10. A large number of reference devices designed from the theory discussed in this paper are in service in military and commercial applications.

Fig. 13 shows a production model of an aircraft alternating voltage regulator in which a static reference such as described in this paper has been employed. The dimensions of the regulator shown are approximately 13 by 6 by 5 inches. The reference is not visible, but it is contained in a package similar to that marked with the arrow in the figure.

A reference used for unattended central station regulator systems[2] has the following characteristics:

Input

105 to 120 volts.
400 to 420 cycles per second.

Output

Constant within ±0.1 per cent.

Conclusions

The theory of operation of a new type of static-reference device has been described. The components involved are inherently rugged and reliable, and the reference should have a long, trouble-free life. The reference device is equally suited for military and commercial applications. Determination of the accuracy of a reference as a function of size, weight, input variations, and power loss has not been completely explored here. Equations can be written for these conditions, but for a given size and power loss, it is usually simpler to design the reference and check the accuracy afterwards. For maximum accuracy, the design of the components is indicated in the theory.

References

1. BIASED RECTIFIERS APPLIED TO MAGNETIC AMPLIFIERS, **Michel Mamon.** *Electrical Manufacturing*, New York, N. Y., Aug. 1953.

2. A NEW REGULATOR AND EXCITATION SYSTEM, **J. T. Carleton, P. O. Bobo, W. F. Horton.** *AIEE Transactions*, vol. 72, pt. III, April 1953, pp. 175–83.

Discussion

R. R. Jackson (National Research Council, Ottawa, Ontario, Canada): This is an excellent description of the use of the saturation flux density of a square-loop material as a reference.

It is interesting to compare the author's device with that described by Milnes and Vernon[1] which uses exactly the same principle. Though some differences exist in the means of compensation of errors attributable to variations in voltage, frequency, and temperature, the voltage reference is derived from the saturating flux density of a high-permeability material.

The following figures are given for the Milnes-Vernon device.

For input: 104 to 127 volts; 370 to 430 cycles per second; 7 volt-amperes.

For output: 8 milliamperes ± 1 per cent into a 60-ohm load during 1,000 hours, for an ambient temperature variation from +15 to +35 C, without temperature compensation.

REFERENCE

1. A SATURABLE-CORE REFERENCE SOURCE FOR USE WITH MAGNETIC AMPLIFIERS, **A. G. Milnes, T. V. Vernon.** *Journal of Scientific Instruments*, London, England, vol. 30, 1953, pp. 135–38.

F. J. Ellert (General Electric Company, Schenectady, N.Y.): Reference circuits, such as those so ably described by the author, engender considerable interest. As far as the past is concerned, the paper lists only two items in the list of references. Since many readers may be interested in some of the historic background, I should like to contribute the following items.

REFERENCES

1. SELBSTTÄTIGE STROM-UND SPANNUNGSREGLER, **W. Geyger.** *Archiv für technisches Messen*, Munich, Germany, J062-7, Nov.; J062-8, Dec. 1934.

2. VOLTAGE REGULATORS USING MAGNETIC SATURATION, **K. J. Way.** *Electronics*, New York, N.Y., July 1937, pp. 14–16.

The Transmission Matrix of N Alike Cascaded Networks

LEO STORCH
NONMEMBER AIEE

Synopsis: The transmission matrix of N alike cascaded networks is derived without the help of advanced concepts or theorems of matrix algebra. Thereafter, a bridge is provided between this more modern approach and the older iterative-parameter theory.

THERE are many applications which suggest the cascade connection of a number of alike networks to achieve a desired over-all characteristic with relatively simple building blocks. While the compact formulation of the transfer function of N cascaded symmetrical networks is probably fairly well known, particularly for image-impedance termination, this does not seem to be the case for a chain of unsymmetrical networks terminated by an arbitrary load impedance. Possibly, this is because the iterative-parameter formulation is unwieldy and somewhat obscure while the more modern publications on this subject are couched in terms of advanced concepts and theorems of matrix algebra.

In this paper the transmission matrix of N alike cascaded networks is developed by means of fairly elementary operations, and a bridge is provided between this modern approach and the older iterative-parameter theory.

Over-all Transmission Matrix of N Alike Cascaded Networks

If the transmission matrix[1] of a certain 2-terminal-pair network is $\|A\|$ (Fig. 1)

Paper 54-513, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted April 30, 1954; made available for printing September 10, 1954.

LEO STORCH is with the Hughes Aircraft Company, Culver City, Calif.

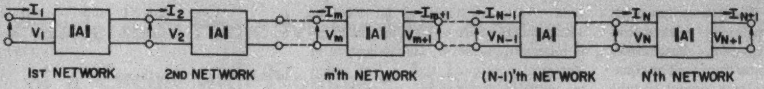


Fig. 1. Cascade connection of N alike networks

$$\left\|\begin{matrix}V_m\\I_m\end{matrix}\right\|=\|A\|\cdot\left\|\begin{matrix}V_{m+1}\\I_{m+1}\end{matrix}\right\|=\left\|\begin{matrix}a, & b\\c, & d\end{matrix}\right\|\cdot\left\|\begin{matrix}V_{m+1}\\I_{m+1}\end{matrix}\right\| \quad (1)$$

then the transmission matrix of a cascade of N such networks is $\|A\|^N$, provided they are all oriented the same way. It would be hard to surpass this statement in compactness, which must be credited to the formulation of the transmission characteristics in terms of matrix algebra. But, if left in this form, it still takes $N-1$ repeated matrix multiplications to express $\|A\|^N$ explicitly in terms of a, b, c, and d. Rather than perform these repeated matrix multiplications, ways can be found to express $\|A\|^N$ directly as a linear combination of $\|A\|$ and the second-order unit matrix $\|I\|$ for any value of N, where the coefficients of these two matrices are ordinary polynomials in $(a+d)$.

Two fine contributions[2,3] concerned with passive networks have appeared on this subject recently. In view of the probable interest in such an equation, even those acquainted with these references may not think it redundant to be presented with a more elementary derivation, which possibly even has the additional merit of arriving at these polynomials by a more direct approach.

Summarizing briefly, Pease[2] uses the Cayley-Hamilton theorem to express $\|A\|^2$ as a linear combination of $\|A\|$ and $\|I\|$. On the basis of this equation, he intuitively assumes a general solution of $\|A\|^N$ in terms of Tchebyscheff polynomials of the second kind, i.e.

$$\frac{\sinh\left(N\cosh^{-1}\dfrac{(a+d)}{2}\right)}{\sinh\left(\cosh^{-1}\dfrac{(a+d)}{2}\right)}$$

and then proves its correctness by induction. Armstrong,[3] on the other hand, resorts to Sylvester's theorem, which allows him to work out the general solution without the help of intuition, but this derivation is probably less pleasing than Pease's in other respects. The general solution for $\|A\|^N$ will be derived here without invoking advanced concepts such as the Cayley-Hamilton theorem, Sylvester's theorem, or even the characteristic equation and the latent roots of a matrix. In a way, this method constitutes an extension of the approach used in a previous paper[4] to derive the transmission properties of N cascaded networks possessing a specific configuration. The point of view adopted in both cases favors the construction of the pertinent polynomials by a simple recursive process rather than the intuitive postulation of a general solution and its subsequent proof of legitimacy by means of induction.

Let $\|A\|$ be multiplied by itself, i.e., squared

$$\|A\|^2=\left\|\begin{matrix}a^2+bc, & b(a+d)\\c(a+d), & d^2+bc\end{matrix}\right\| \quad (2)$$

Since $(a+d)$ is already a factor of the elements in one diagonal, it can be factored from each element of the matrix after adding and subtracting ad in the other diagonal and splitting off a diagonal matrix

$$\|A\|^2=\left\|\begin{matrix}a(a+d), & b(a+d)\\c(a+d), & d(a+d)\end{matrix}\right\|-\left\|\begin{matrix}ad-bc, & 0\\0, & ad-bc\end{matrix}\right\|$$

or

$$\|A\|^2=(a+d)\|A\|-(ad-bc)\|I\| \quad (3)$$

where $\|I\|$ is the second-order unit matrix. When the network is bilateral, i.e., passive and does not contain gyrators, $ad-bc=1$ and

$$\|A\|^2=(a+d)\|A\|-\|I\| \quad (4)$$

Since $\|A\|^u\cdot\|I\|^v=\|I\|^v\cdot\|A\|^u=\|A\|^u$ for any integers u and v, the powers of $\|A\|$ may be expanded by the ordinary binomial theorem $(y=a+d)$

$$\|A\|^{2M}=[y\|A\|-\|I\|]^M=\sum_{s=0,1,2,\ldots}(-1)^{M-s}\binom{M}{s}y^s\|A\|^s \quad (5A)$$

and

$$\|A\|^{2M+1}=\|A\|\cdot\|A\|^{2M}=\sum_{s=0,1,2,\ldots}(-1)^{M-s}\binom{M}{s}y^s\|A\|^{s+1} \quad (5B)$$

Inasmuch as each term $\|A\|^s$, $s>1$, may be re-expanded and simplified by equation 4, it follows that, for any positive integer N, $\|A\|^N$ may be written as follows

$$\|A\|^N=P_{(y)N}\|A\|-R_{(y)N}\|I\| \quad (6)$$

Both $P_{(y)N}$ and $R_{(y)N}$ are ordinary scalar polynomials in y and are associated with the N'th power of $\|A\|$. Moreover

$$\|A\|^{N+1}=\|A\|\cdot\|A\|^N=P_{(y)N}\|A\|^2-R_{(y)N}\|A\|=(yP_{(y)N}-R_{(y)N})\|A\|-P_{(y)N}\|I\| \quad (7)$$

by means of equation 4. Since $\|A\|^{N+1}$ is also equal to $P_{(y)N+1}\|A\|-R_{(y)N+1}\|I\|$, it follows that

$$R_{(y)N+1}=P_{(y)N} \quad (8)$$

and

$$P_{(y)N+1}=yP_{(y)N}-P_{(y)N-1} \quad (9)$$

The over-all transmission matrix of N alike cascaded bilateral networks is therefore

$$\|A\|^N=P_{(y)N}\|A\|-P_{(y)N-1}\|I\| \quad (10)$$

and the scalar polynomial $P_{(y)M}$ obeys the recursion equation $(M\geq 3)$

$$P_{(y)M}=yP_{(y)M-1}-P_{(y)M-2} \quad (11)$$

The recursion process is started off by $P_{(y)1}$ and $P_{(y)2}$

$$P_1=1$$

$$P_2=y$$

which are derived from $\|A\|^1=\|A\|$ and $\|A\|^2=y\|A\|-\|I\|$

The higher order polynomials are generated easily by recursion

$$P_3=y^2-1$$

$$P_4=y^3-2y$$

$$P_5=y^4-3y^2+1$$

$$P_6=y^5-4y^3+3y$$

$$P_7=y^6-5y^4+6y^2-1$$

$$P_8=y^7-6y^5+10y^3-4y$$

or

$$P_8=\binom{7}{0}y^7-\binom{6}{1}y^5+\binom{5}{2}y^3-\binom{4}{3}y$$

$$\vdots$$

$$P_{(y)M}=\binom{M-1}{0}y^{M-1}-\binom{M-2}{1}y^{M-3}+\binom{M-3}{2}y^{M-5}\ldots$$

Since binomial expansions determine the polynomials $P_{(y)M}$, it is not too far-fetched to expect binomial coefficients to appear in them. The pattern established by the columns in the tabulation of $P_{(y)1}$ through $P_{(y)8}$ is easily recognized[5] and generalized for any M

$$P_{(y)M}=\sum_{s=0,1,2,3,\ldots}(-1)^s\binom{M-1-s}{s}y^{M-1-2s} \quad (12)$$

For the most general linear 2-terminal-pair network, $ad-bc$ is not equal to 1. Let the value of this determinant be designated by δ. Then the polynomials P_N and R_N are functions of δ as well as of y, since the expansions derive from equation 3 rather than from 4. The equations corresponding to 8, 9, and 12 take the form

$$R_{N+1}=\delta P_N \tag{8A}$$

$$P_{N+1}=yP_N-\delta P_{N-1} \tag{9A}$$

where $P_1=1$, $P_2=y$ as before

$$P_M=\sum_{s=0,1,2,\ldots}(-\delta)^s\binom{M-1-s}{s}y^{M-1-2s} \tag{12A}$$

The polynomials, equation 12, have received the name "Tchebyscheff polynomials of the second kind" in the mathematical literature, since they can also be expressed[6] as

$$\frac{\sinh\left[M\cosh^{-1}\dfrac{y}{2}\right]}{\sinh\left[\cosh^{-1}\dfrac{y}{2}\right]}$$

and have certain properties which are interesting, but fairly immaterial as far as this particular application is concerned. The equivalence is mentioned here primarily since it may be useful at times in numerical calculations. In developing $\|A\|^N$, however, it appears preferable to avoid the transcendental functions entirely and to obtain the solution by simple rational operations as has been done here. This conclusion is corroborated also by the next section, which is devoted to the solution in terms of iterative parameters, i.e., hyperbolic functions.

Iterative Parameters and the Over-all Transmission Matrix of N Alike Cascaded Networks

Many readers may be wondering why several contributions have been appearing which deal with the transmission characteristics of a chain of N alike bilateral networks in terms of those of a single of these networks. After all, this is the job assigned in classical network theory to the iterative parameters of a network. The correct answer is probably that iterative parameters are very unwieldy and uninviting. Although discussed in textbooks for more than a quarter of a century, they have apparently never acquired a secure foothold. Johnson[7] states, as far back as 1924, that they are of little interest.[8] It is intriguing, however, and indicative of historical continuity that a few added manipulations transform the iterative-parameter formulation of the over-all transmission characteristics into the $\|A\|^N$ of equation 10.

It is a telling point that, although the iterative impedances Z_{K_1} and Z_{K_2} as well as the propagation constant Γ are discussed in many places, explicit expressions for the general circuit parameters a, b, c, and d in terms of the iterative parameters are usually not included. However, they may be obtained without difficulty from the equations expressing Z_{K_1}, Z_{K_2}, and Γ as functions of a, b, c, and d. They may also be obtained, for instance, from the equivalent T network[9] with the arms expressed in terms of Z_{K_1}, Z_{K_2}, and Γ as given by Johnson[10]

$$a=\cosh\Gamma+\frac{Z_{K_1}-Z_{K_2}}{Z_{K_1}+Z_{K_2}}\sinh\Gamma \tag{13}$$

$$c=\frac{2\sinh\Gamma}{Z_{K_1}+Z_{K_2}} \tag{14}$$

$$d=\cosh\Gamma-\frac{Z_{K_1}-Z_{K_2}}{Z_{K_1}+Z_{K_2}}\sinh\Gamma \tag{15}$$

$$b=\frac{ad-1}{c}=\frac{2Z_{K_1}Z_{K_2}}{Z_{K_1}+Z_{K_2}}\sinh\Gamma \tag{16}$$

By physical reasoning[11] it is established that the general circuit parameters a_N, b_N, c_N, and d_N for N cascaded networks of the same type are obtained by replacing Γ by $N\Gamma$ in equations 13 through 16

$$a_N=\cosh N\Gamma+\frac{Z_{K_1}-Z_{K_2}}{Z_{K_1}+Z_{K_2}}\sinh N\Gamma$$

$$b_N=\frac{2Z_{K_1}Z_{K_2}}{Z_{K_1}+Z_{K_2}}\sinh N\Gamma$$

$$c_N=\frac{2\sinh N\Gamma}{Z_{K_1}+Z_{K_2}}$$

$$d_N=\cosh N\Gamma-\frac{Z_{K_1}-Z_{K_2}}{Z_{K_1}+Z_{K_2}}\sinh N\Gamma$$

Rather than stop at this point, as is commonly done, let Z_{K_1} and Z_{K_2} be eliminated immediately from b_N and c_N by writing the following

$$b_N=b\frac{\sinh N\Gamma}{\sinh\Gamma} \tag{17}$$

$$c_N=c\frac{\sinh N\Gamma}{\sinh\Gamma} \tag{18}$$

In view of the identity

$$\cosh N\Gamma=\cosh\Gamma\frac{\sinh N\Gamma}{\sinh\Gamma}-\frac{\sinh(N-1)\Gamma}{\sinh\Gamma} \tag{19}$$

it is easy to see, furthermore, that

$$a_N=a\frac{\sinh N\Gamma}{\sinh\Gamma}-\frac{\sinh(N-1)\Gamma}{\sinh\Gamma} \tag{20}$$

$$d_N=d\frac{\sinh N\Gamma}{\sinh\Gamma}-\frac{\sinh(N-1)\Gamma}{\sinh\Gamma} \tag{21}$$

Since[12]

$\cosh\Gamma=(a+d)/2=y/2$

$(\sinh M\Gamma)/(\sinh\Gamma)$ may be written as a polynomial[6] in y

$$\begin{aligned}\frac{\sinh M\Gamma}{\sinh\Gamma}&=2^{M-1}\cosh^{M-1}\Gamma-\binom{M-2}{1}\times\\&\quad 2^{M-3}\cosh^{M-3}\Gamma+\binom{M-3}{2}2^{M-5}\times\\&\quad\cosh^{M-5}\Gamma-\binom{M-4}{3}2^{M-7}\cosh^{M-7}\Gamma\ldots\\&=\sum_{s=0,1,2,\ldots}(-1)^s\binom{M-1-s}{s}y^{M-1-2s}=\\&\qquad P_{(y)M}\end{aligned} \tag{22}$$

Therefore, there is complete agreement with equation 10, as can be seen by collecting equations 20, 17, 18, and 21 in the form of a matrix after replacing $(\sinh M\Gamma)/(\sinh\Gamma)$ by $P_{(y)M}$ because of the identity between equations 12 and 22

$$\|A\|^N=\begin{Vmatrix}a_N, & b_N\\ c_N, & d_N\end{Vmatrix}=P_{(y)N}\|A\|-P_{(y)N-1}\|I\|$$

References

1. **Communications Networks, E. A. Guillemin.** John Wiley & Sons, Inc., New York, N. Y., vol. II, chap. IV, 1935.

2. **The Iterated Network and Its Application to Differentiators, M. C. Pease.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 40, no. 6, June 1952, pp. 709–11.

3. **Note on the Iterated Network and Its Application to Differentiators, H. L. Armstrong.** *Ibid.*, vol. 41, no. 5, May 1953, p. 667.

4. **The Multisection RC Filter Network Problem, L. Storch.** *Ibid.*, vol. 39, no. 11, Nov. 1951, pp. 1456–58.

5. **Smithsonian Mathematical Formulae and Tables of Elliptic Functions.** *Smithsonian Institution*, Washington, D. C., 1922, p. 20.

6. *Ibid.*, p. 67, formula no. 3.175.

7. **Transmission Circuits for Telephonic Communication, K. S. Johnson.** D. van Nostrand Company, Inc., New York, N. Y., 1924, chap. XI.

8. *Ibid.*, pp. 134, 137, 139, 141.

9. **Guillemin,** *loc. cit.* Substitute z_{11}, z_{12}, z_{22} from equation 339 on p. 156 in equations 295 on p. 138.

10. **Johnson,** *loc. cit.*, Fig. 16, p. 135. (Γ corresponds to P in Johnson's notation).

11. *Ibid.*, p. 124.

12. **Guillemin,** *loc. cit.*, equation 363, p. 164.

Acceleration Plane Method for Analysis of a Circuit with Nonlinear Inductance and Nonlinear Capacitance

Y. H. KU
FELLOW AIEE

Synopsis: This paper analyses the transient behavior of a circuit with nonlinear resistance, inductance, and capacitance by the acceleration plane method developed recently by the author.[1] The method as extended here is able to solve a nonlinear differential equation of the following type: $\phi(\dot{x})\ddot{x} + f(\dot{x}, x) + f_1(x) = F(t)$. Five examples are given, with $F(t)$ equal to a constant or a time function. Either $\phi(\dot{x})$, $f(\dot{x},x)$, or $f_1(x)$ may be gven as a graph plotted from experimental data. The forcing function $F(t)$ may be a sine wave or any time function. The method owes its simplicity to a close relation with physical boundary conditions. In one example, the phenomenon of ferroresonance is obtained. Also the possibility of applying the same method to a circuit with $R(t)$, $L(t)$, and $C(t)$ is indicated.

IN a recent paper[1] the author developed an acceleration plane method by which a nonlinear differential equation of the following type can be solved

$$\ddot{x}+f(\dot{x}, x)+f_1(x)=F(t) \tag{1}$$

where

x = displacement or charge
q, $\dot{x}$ = velocity $\mathbf{v}$ or current i
$\ddot{x}$ = acceleration g or di/dt

Since acceleration is equal to $d\mathbf{v}/dt$, where $\mathbf{v}$ represents velocity

$$\frac{d\mathbf{v}}{dt}=F(t)-f(\mathbf{v}, x)-f_1(x)=G(t, \mathbf{v}, x) \tag{2}$$

Dividing equation 2 by $\mathbf{v}=\dot{x}$

$$\frac{d\mathbf{v}}{dx}=\frac{G(t, \mathbf{v}, x)}{\mathbf{v}}=\frac{F(t)-f(\mathbf{v}, x)-f_1(x)}{\mathbf{v}} \tag{3}$$

Notice that $G(t, \mathbf{v}, x)$ is an acceleration function of the most general type involving time function, velocity, and displacement in a nonlinear way. According to equation 3, a method can be developed of constructing the proper slope $d\mathbf{v}/dx$ in the velocity plane from the ratio of acceleration to velocity.

In electrical terms

$$\frac{di}{dt}=E(t)-f(i, q)-f_1(q)=G(t, i, q) \tag{2A}$$

$$\frac{di}{dq}=\frac{E(t)-f(i, q)-f_1(q)}{i}=\frac{G(t, i, q)}{i} \tag{3A}$$

where $E(t)$ = voltage as a time function. Notice that a circuit with nonlinear resistance and nonlinear capacitance can be treated by the method outlined in reference 1.

Circuit with Nonlinear Inductance

For a circuit with nonlinear inductance and constant capacitance, the differential equation takes the following form

$$\dot{\phi}+Ri+\frac{q}{C}=E(t) \tag{4}$$

where

ϕ represents the flux which is a complicated function of i depending on the curve of magnetic induction versus magnetic intensity (B-H curve) and reduces to Li for

L = constant inductance
R = resistance
C = capacitance

Keller[2] analyzed this problem by expressing i as a function of the flux density B in a power series and checked Butler and Concordia's results[3] obtained from differential analyzer studies. Pipes,[4] on the other hand, used ϕ as a function of current i in a power series and showed an example by taking the first two terms of the series.

The acceleration plane method as presented in this paper can be applied to the actual case where $\phi(i)$ is given from the experimental data, say, the B-H curve for a certain material. In the following discussion, $\phi(i)$ means flux ϕ as a function of i as plotted in the form of a graph (similar to the B-H curve), not necessarily as an explicit function or a known power series fitting the experimental data approximately.

Since ϕ is given as a function of i in the form of a graph similar to the B-H curve (other physical problems may have similar data available) the slope of $d\phi/di$ must first be obtained and a plot made, see Fig. 1. This slope curve or graph will be denoted by ϕ'. Thus equation 4 becomes

$$\left(\frac{d\phi}{di}\right)\frac{di}{dt}+Ri+\frac{q}{C}=E(t) \tag{5}$$

or

$$\phi'\frac{di}{dt}+Ri+\frac{q}{C}=E(t) \tag{5A}$$

Notice that ϕ' reduces to L in the constant inductance case where $\phi=Li$. If there is an additional constant inductance L_1 besides the nonlinear inductor, ϕ' can be replaced by $\phi'+L_1$.

Choosing $\phi=L \tanh i$ as an illustration and dividing both members of equation 5 by L

$$(\operatorname{sech}^2 i)\frac{di}{dt}+\frac{R}{L}i+\frac{q}{LC}=\frac{E(t)}{L} \tag{5B}$$

Notice that either equation 5(A) or 5(B) belongs to the general type of nonlinear differential equation

$$\phi(\dot{x})\ddot{x}+f(\dot{x}, x)+f_1(x)=F(t) \tag{6}$$

Since equation 5(B) differs from the standard form given in equation 2(A) by the factor $\operatorname{sech}^2 i=1/\cosh^2 i$, every term in equation 5(B) is premultiplied by $\cosh^2 i$ so that $G(t, i, q)$ in equation 2(*A*) will be further generalized as an acceleration function. Notice that $G(t, i, q)$ will now include a term $\cosh^2 i[E(t)/L]$.

Thus, in general

$$\frac{di}{dt}=G(t, i, q)=(1/\phi')[E(t)-f(i, q)-f_1(q)] \tag{7}$$

In this particular example, $\phi=L \tanh i$, and $1/\phi'=(\cosh^2 i)/L$. So

$$\frac{di}{dt}=\frac{\cosh^2 i}{L}[E(t)-f(i, q)-f_1(q)] \tag{7A}$$

If, as before, $f(i, q)=Ri$ and $f_1(q)=q/C$, and if $a=R/2L$ and $\omega_0^2=1/LC$, equation 7(A) becomes

$$\frac{di}{dt}=\cosh^2 i\left(\frac{E(t)}{L}-2ai-\omega_0^2 q\right) \tag{7B}$$

Since equation 2(A) has been generalized to equation 7, equation 3(A) can be similarly generalized to

$$\frac{di}{dq}=\frac{(1/\phi')[E(t)-f(i, q)-f_1(q)]}{i}=\frac{G(t, i, q)}{i} \tag{8}$$

Paper 54-514, recommended by the AIEE Basic Sciences Committee, and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11-15, 1954. Manuscript submitted March 16, 1953; made available for printing August 24, 1954.

Y. H. Ku is with the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pa.

The author wishes to acknowledge his indebtedness to Vice President C. C. Chambers, Drs. J. G. Brainerd, S. R. Warren, Jr., and C. N. Weygandt for encouragement and suggestions.

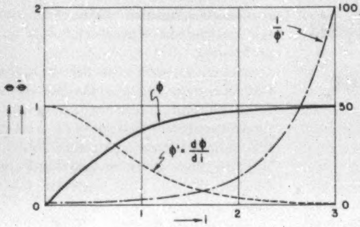


Fig. 1. Flux-current graph and slope curves

$$\phi = \tanh i$$

$$\phi' = \frac{1}{\cosh^2 i}$$

$$\frac{1}{\phi'} = \cosh^2 i$$

In terms of velocity and displacement, this becomes

$$\frac{d\mathbf{v}}{dx}=\frac{(1/\phi')[F(t)-f(\mathbf{v},x)-f_1(x)]}{\mathbf{v}}=\frac{G(t,\mathbf{v},x)}{\mathbf{v}} \tag{8A}$$

where $\phi(\mathbf{v})$ is given in equation 6, and $\phi' = d\phi/d\mathbf{v}$.

Illustrative Example I

The solution of a linear circuit with R, L, and C all constants is presented here to serve as a basis of comparison. For $\phi = Li$, where L=constant inductance, equation 7(B) is simplified to

$$\frac{di}{dt}=\frac{E(t)}{L}-2ai-\omega_0^2 q \tag{9}$$

where

$a = R/2L$
$\omega_0^2 = 1/LC$

Noting $i = \dot{q}$ and $di/dt = \ddot{q}$ and putting $t' = \omega_0 t$, equation 9 can be normalized to

$$\frac{d^2q}{dt'^2}=\frac{E(t')}{L/LC}-k\frac{dq}{dt'}-q \tag{10}$$

where

$t' = t/\sqrt{LC}$
$k = R\sqrt{C/L}$

For a numerical example, $R=0.2$, $L=0.4$, $1/C=0.4$, and $E=1.0$ such that $LC=1$, $t'=t$, and $k=0.5$.

Equation 10 then becomes

$$\frac{di}{dt}+0.5i+q=2.5 \tag{11}$$

which gives

$$\frac{di}{dq}=\frac{2.5-0.5i-q}{i}=\frac{G(i,q)}{i}=\frac{g}{i} \tag{11A}$$

where g represents a simplified form of the general acceleration function $G(t, \mathbf{v}, x)$ or $G(t, i, q)$ for $F(t)$=constant.

The $\mathbf{v}-x$ or $i-q$ plot can be constructed according to the acceleration plane method as follows:

First, acceleration versus displacement plots or $g-x$ plots for constant values of $\mathbf{v}$ are made. Since $g=2.5-0.5\ \mathbf{v}-x$ according to equation 11(A), for $\mathbf{v}=i=0$

$$g=G(0, q)=2.5-0.5(0)-q=2.5-q$$

As shown in Fig. 2, with acceleration g or di/dt as ordinate and displacement x or charge q as abscissa, the $g-x$ curve (or $di/dt-q$ plot) for $\mathbf{v}=i=0$ is a straight line passing through the following two points: $q=0$, $g=2.5$, and $q=2.5$, $g=0$. In fact, these two points on the $\mathbf{v}=i=0$ line plotted on the $g-x$ plane (or simply acceleration plane) represent two important boundary conditions of the physical problem. When a voltage is suddenly applied to a R, L, C series circuit at $t=0$ $q=0$, $i=0$, and $di/dt=g=E/L=1/0.4=2.5$. As time goes on, the transient components die out, and the final steady-state condition gives $g=di/dt=0$, $i=0$ and $q = EC = 1(1/0.4) = 2.5$

Fov $\mathbf{v}=i=1$, 2, or 3,

$$g=G(1, q)=2.5-0.5(1)-q=2.0-q$$

$$g=G(2, q)=2.5-0.5(2)-q=1.5-q$$

$$g=G(3, q)=2.5-0.5(3)-q=1.0-q$$

Similarly, for $\mathbf{v}=i=\ -1$

$$g=G(-1, q)=2.5-0.5(-1)-q=3.0-q$$

From Fig. 2 it is seen that these equi-velocity or equicurrent curves are parallel straight lines on the acceleration plane plot.

Although equation 11(A) gives the correct slope everywhere on the $\mathbf{v}-x$ or $i-q$ plane (or simply velocity plane), this method differs from the isocline method in that there is no need of plotting all the possible slopes and then finding the $\mathbf{v}-x$ curve or trajectory. This method starts from a given boundary condition on the $\mathbf{v}-x$ plot, say, $\mathbf{v}=i=0$ and $x=q=0$, and assumes an initial slope starting from the origin. Suppose an almost vertical slope

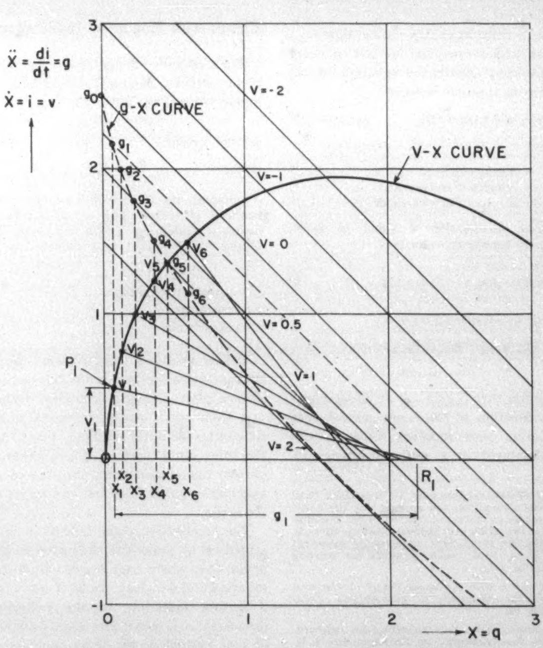


Fig. 2. Construction of v−x trajectory by the acceleration plane method

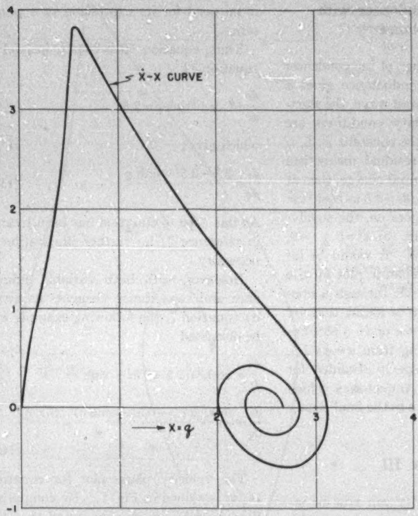


Fig. 3. Velocity trajectory or v − x curve

$\phi(\dot{x})\ddot{x}+k\dot{x}+x=$ constant

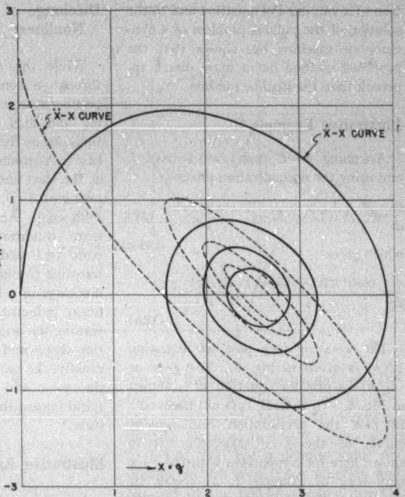


Fig. 5. Acceleration and velocity trajectories

$\ddot{x}+0.5\dot{x}+x=2.5$

is assumed in the beginning as it is known that (di/dq) at $i=0$ and $q=0$ is equal to $g/i=2.5/0=\infty$. As a steep positive slope is taken upward, there will be a comparatively large increment of Δi for a corresponding small increment of Δq to the right of the origin. A suitable pair of values of $i_1=0+\Delta i$ and $q_1=0+\Delta q$ is chosen along the assumed slope line. From this chosen value of q_1, a vertical line is drawn upward to locate a proper value of g_1 from the family of $i=$constant curves described in the first step of the plot. This point g_1 is shown in Fig. 2, corresponding to $q=q_1$ and $i=i_1$. Knowing g_1, a distance to the right of $q_1=x_1$ is marked off on the horizontal axis and a point R_1 is located. A straight line joining R_1 and P_1 is drawn, which is determined by $v_1=i_1$ and $x_1=q_1$. A line drawn perpendicular to the R_1P_1 line will determine the proper slope at P_1. If this slope checks with the assumed slope, P_1 will be a proper point on the $v-x$ or $i-q$ trajectory. It is necessary to take small steps and check the slopes carefully in the beginning. As the construction gets properly started, the steps as shown in Fig. 2 are sufficient to make further progress in plotting the $v-x$ or $i-q$ curve or trajectory. The $g-x$ trajectory is also shown.

Application[1] of the foregoing construc-

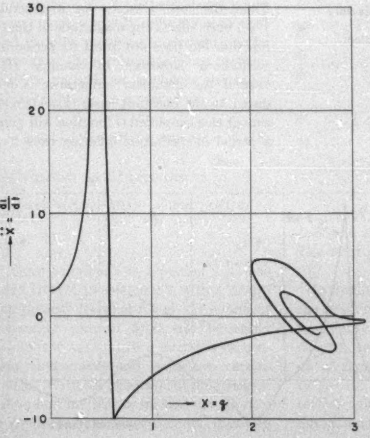
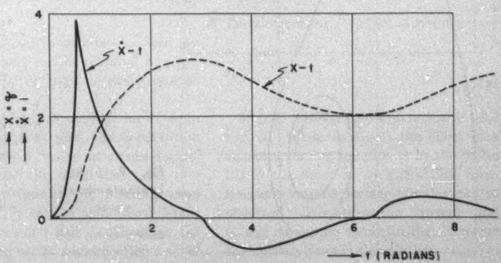


Fig. 4 (left). Acceleration trajectory or g − x curve

$\phi(\dot{x})\ddot{x}+k\dot{x}+x=$ constant

Fig. 6 (below). Instantaneous waves

$\phi(\dot{x})\ddot{x}+k\dot{x}+x=$ constant



tion to a van der Pol equation and to the solution of the pullout problem of a synchronous machine has shown that the proposed method has a more direct approach than the isocline method.

Illustrative Example II

Assuming $\phi=L \tanh i=0.4 \tanh i$, and using the normalization process

$$\frac{di}{dt}=\cosh^2 i(2.5-0.5i-q) \tag{12}$$

which gives

$$\frac{di}{dq}=\frac{\cosh^2 i(2.5-0.5i-q)}{i}=\frac{G(i, q)}{i}=\frac{g}{i} \tag{12A}$$

The $\mathbf{v}-x$ or $i-q$ plot for equation 12(A) is given in Fig. 3. The $g-x$ or $(di/dt)-q$ plot for equation 12 is shown in Fig. 4. In general, it is not necessary to plot the acceleration displacement curve, or the $g-x$ trajectory. It is shown here for comparison with the $g-x$ trajectory of example I, as shown in Fig. 5. In Fig. 5, di/dt starts from 2.5, decreases to negative values, and then spirals towards 0 at $q=2.5$. In Fig. 4 di/dt starts from 2.5, increases to as high a value as 63, decreases rapidly to -10, and then spirals to 0 at $q=2.5$ as shown. The $i-t$ and $q-t$ waves are shown in Fig. 6. The general shape of the current wave is familiar to those who have experimented with oscillographic records of transformer inrush currents.

Discharge of a Condenser with Nonlinear Inductance

While the charging of a condenser through a nonlinear inductance gives a peculiar transient current wave, the starting and final boundary conditions are the same as those in the constant R, L, C case. Assuming no residual magnetism at the start and no initial charge, then at $t=0, q=0, i=0$, and $di/dt=2.5$ as $\cosh^2 i=1$ for $i=0$. As time goes on, the steady-state condition gives $di/dt=0$, $i=0$, $\cosh^2 i=1$, and $q=2.5$. It would be interesting to make a separate plot for the discharge of a condenser through a nonlinear inductance. It is found that by turning the original $\mathbf{v}-x$ or $i-q$ plot upside down and starting from $x=q=2.5$, exactly the same shape is obtained for the $\mathbf{v}-x$ and $g-x$ trajectories which spiral toward the origin in the final steady state.[5]

Illustrative Example III

A further nonlinearity can now be introduced in the condenser element by using $f_1(q)$ instead of q/C. Again, it may be made clear at the beginning that $f_1(q)$ may be any function of q obtained from experimental data. To simplify the example which illustrates the general method, however, it will be assumed that $f_1(q)=(1/C) \sinh q$, as outlined by Pipes.[6] Again, the whole function $\sinh q$ or any similar graph will be considered and there is no need for its expression in a power series.

Thus, equation 10 becomes, instead of equation 11

$$\frac{di}{dt}+0.5i+ \sinh q=2.5 \tag{13}$$

which gives

$$\frac{di}{dq}=\frac{2.5-0.5i-\sinh q}{i} \tag{13A}$$

As this type of equation has been treated in reference 1, no further illustration is necessary.

However, with both variable inductance and capacitance elements as given by equation 7, the following example will be discussed

$$\frac{di}{dt}= \cosh^2 i(2.5-0.5i- \sinh q) \tag{14}$$

$$\frac{di}{dq}=\frac{\cosh^2 i(2.5-0.5i- \sinh q)}{i}=\frac{G(i, q)}{i} \tag{14A}$$

The velocity plane plot for equation 14(A) is shown in Fig. 7. By comparing Figs. 3 and 7, it may be noted that the $\mathbf{v}-x$ curve or the $i-q$ curve for example III has the general shape of that for example II. However, as in the final steady state, $\sinh q=2.5$ gives $q=1.65$, the focal point is decreased from 2.5 in example II (Fig. 3) to 1.65 in example III (Fig. 7). Also, the natural frequency of oscillation in example III is no longer determined by $1/\sqrt{LC}$ and also differs from that obtained in example II.

The instantaneous current (or velocity) and charge (or displacement) values are plotted in Fig. 8. It is noted that in Fig. 6, the period of one complete cycle is about 2π, as LC is taken as 1. Thus, the $\cosh^2 i$ factor (or, in general, $1/\phi'$) only affects the magnitude of the G function but does not affect its period of oscillation. However, in example III, even if the nonlinear inductance is reduced to the constant case, $(2.5-0.5i-\sinh q)$ as a simplified G function will give a period of oscillation differing from 2π.

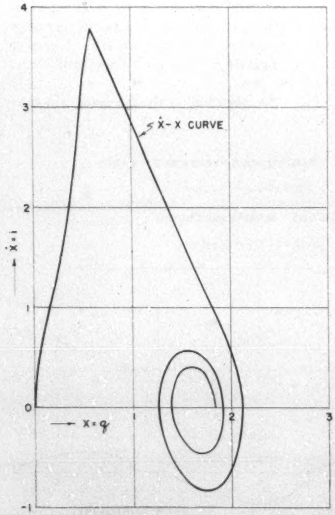


Fig. 7 (left). Velocity trajectory

$\phi(\dot{x})\ddot{x}+k\dot{x}+f_1(x)=\text{constant}$

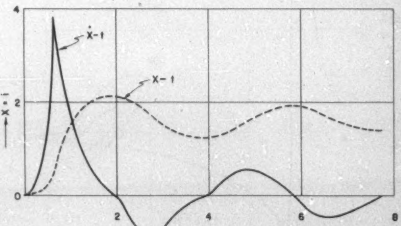


Fig. 8 (right). Instantaneous waves

$\phi(\dot{x})\ddot{x}+k\dot{x}+f_1(x)=\text{constant}$

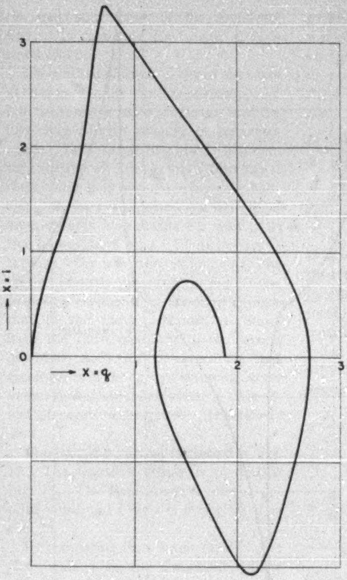


Fig. 9 (left). Velocity trajectory

$\phi(\dot{x})\ddot{x}+k\dot{x}+x = E_1(t)$

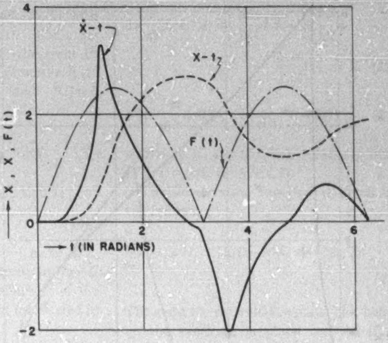


Fig. 10 (right). Instantaneous waves

$\phi(\dot{x})\ddot{x}+k\dot{x}+x = E_1(t)$

As shown in Fig. 8, the period of one complete cycle is found to be 4 radians instead of 2π radians with a ratio of $2/\pi$ for the two cases. In general, it can be said that, with an additional nonlinear element as a function of x or q, the period of oscillation will be changed depending on the function $f_1(x)$ or $f_1(q)$.

Effect of Variable Resistance

In example III it has been shown how a circuit with variable inductance element and variable capacitance element can be treated by the acceleration plane method. With a variable or nonlinear resistance equation 14 can be modified as

$$\frac{di}{dt} = \cosh^2 i[E/L - kf(i) - \sinh q] \qquad (15)$$

and equation 14(A) becomes

$$\frac{di}{dq} = \frac{\cosh^2 i[E/L - kf(i) - \sinh q]}{i} = \frac{G(i, q)}{i} \qquad (15A)$$

Since equation 1 or equation 2(A) has been treated in reference 1 with a damping element $f(\mathbf{v}, x)$ or $f(i, q)$, the nonlinear resistance element $kf(i)$ can be easily taken into account in a circuit with constant inductance and nonlinear capacitance. With the extension of the acceleration plane method to include the effect of a nonlinear inductance element, it is seen that equation 15(A) can be solved by the same procedure used in solving equation 14(A). Thus the method is, in general, applicable to a circuit with nonlinear resistance, nonlinear inductance, and nonlinear capacitance.

Illustrative Example IV

A problem with nonlinear R, L, C elements when a forcing function $F(t)$ or $E(t)$ is suddenly applied will now be solved. From equations 7 and 8

$$\frac{di}{dt} = G(t, i, q) = \cosh^2 i[E(t)/L - kf(i) - \sinh q] \qquad (16)$$

$$\frac{di}{dq} = \frac{G(t, i, q)}{i} = \frac{\cosh^2 i[E(t)/L - kf(i) - \sinh q]}{i} \qquad (16A)$$

In general, $\cosh^2 i$ can be replaced by L/ϕ' and $\sinh q$ can be replaced by $f_1(q)$. The term $E(t)/L$ may be designated as $E_1(t)$.

In finding the proper slope for di/dq in the velocity plane plot, the acceleration function $G(t, i, q)$, as given by equation 16, depends upon the particular value of t at a given instant. If starting from $t=0$, $G(0, i, q)$ can be found for different sets of values of i and q. However, for any assumed set of i and q after a small interval from the initial instant $t=0$, the time interval Δt will be given by $\Delta t = (q_1 - q_0)/(\text{av } i)$, or $\Delta q = i_{av}(\Delta t)$, where i_{av} may be taken as $(i_1+i_0)/2$. Now since $t_1 = t_0 + \Delta t$, $E(t)$ has a definite value according to the given time function and $G(t, i, q)$ at the end of the first interval can be found. It is necessary to check the slope di/dq for a set of values i,q,t such that $di/dq = G(t, i, q)/i$.

To illustrate the method with a given time function, it is sufficient to consider a circuit with nonlinear inductance and make the resistance and the capacitance elements constant. Once the method is understood, the extension to nonlinear R and C elements is rather straightforward.

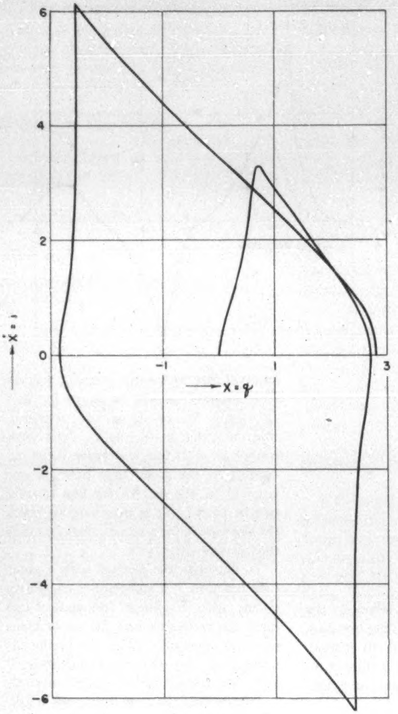
Assume that the time function is a full-wave rectified sine wave with all negative loops changed to positive loops. Again, this assumption is made for the sake of illustration. Any other time function may be treated in a similar manner. Fig. 9 gives the velocity displacement or current charge plot when a circuit with nonlinear inductance and constant resistance and capacitance is subjected to the aforementioned forcing function $F(t)$ or $E(t)$. The generalized acceleration function is given by

$$\frac{di}{dt} = G(t, i, q) = \cosh^2 i[E_1(t) - 0.5i - q] \qquad (17)$$

The slope of the $i-q$ plot is given by

$$\frac{di}{dq} = \frac{\cosh^2 i[E_1(t) - 0.5i - q]}{i} = \frac{G(t, i, q)}{i} \qquad (17A)$$

It is of interest to compare Fig. 9 with Fig. 3. While in Fig. 3 the effect of a constant forcing function is as shown, in Fig. 9 the effect of the given time function as a variable forcing function can be observed. The transient current and charge curves are shown in Fig. 10 where the voltage wave is also plotted.



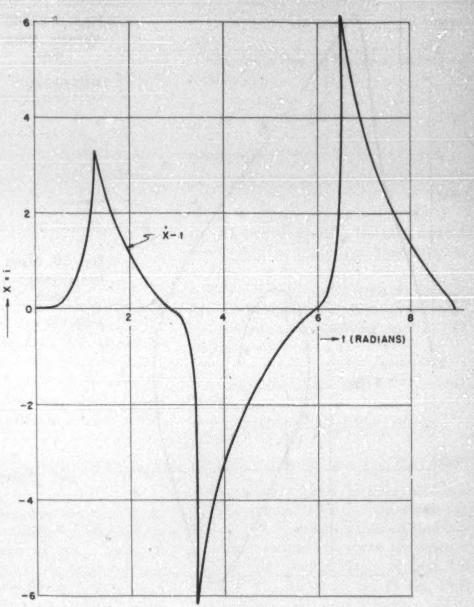


Fig. 11 (left). Velocity trajectory

$\phi(\dot{x})\ddot{x}+k\dot{x}+x=E_1 \sin t$

Fig. 12 (above). Instantaneous current

$\phi(\dot{x})\ddot{x}+k\dot{x}+x=E_1 \sin t$

Effect of Sinusoidal Time Function

The effect of the sudden application of a sinusoidal time function to a nonlinear circuit is of interest both from the theoretical and the practical point of view. In the solution of linear differential equations, transformation calculus has developed a theorem for obtaining an a-c solution when the d-c transient response is known. In fact, by means of superposition integrals, explicit expressions for the transient current caused by an arbitrary electromotive force can be derived. In the nonlinear circuit, however, superposition principle cannot be readily applied. Even in time-varying parameter circuits, special care must be taken in the application of the superposition principle.[7]

For a nonlinear inductive circuit, the phenomenon of ferroresonance has been observed for some time. Both experimental and differential analyzer studies have been reported.[8] This phenomenon comes under the general analytical description of "negative damping" which is found in electric circuits and machinery transients.[8-10] This physical phenomenon can be studied by the nonlinear analysis in the following example.

Illustrative Example V

Assuming the time function to be sin t, where t is in radians, equation 17 becomes

$$\frac{di}{dt}=\cosh^2 i[E_1 \sin t-0.5i-q]=G(t, i, q) \tag{18}$$

where $E_1=E/L$ and E represents the maximum value of the sine wave.

The slope of the $i-q$ plot or the $\mathbf{v}-x$ plot is given by

$$\frac{di}{dq}=\frac{\cosh^2 i(2.5 \sin t-0.5i-q)}{i}=\frac{G(t, i, q)}{i} \tag{18A}$$

Notice that in both examples IV and V,

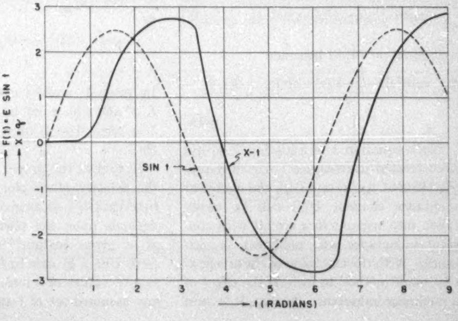


Fig. 13. Instantaneous waves

$\phi(\dot{x})\ddot{x}+kx+x = E_1 \sin t$

the maximum value of the sinusoidal wave was taken as $E=1$, and E/L is equal to 2.5.

The method of solution is just the same as that used in the solution of example IV but the result is more interesting and the plot more difficult because of ferroresonance. It is found that the $i-q$ trajectory (or the $v-x$ trajectory) describes a rather large loop during the negative half-cycle of sin t and keeps on describing another large loop during the next positive half-cycle of sin t. While in Figs. 3 through 9 the trajectory is limited to the first and fourth quadrants, in Fig. 11 the trajectory swings from the fourth quadrant to the third quadrant, continues from the third quadrant to the second quadrant, and then re-enters the first quadrant with the current assuming larger values than when it first started. The effect of negative damping is thus clearly seen.

The transient current is plotted in Fig. 12. The transient charge is plotted in Fig. 13. The time function for the applied voltage, 2.5 sin t is also plotted in Fig. 13.

By comparing Figs. 9 and 11, it is seen that, while a regular sinusoidal voltage gives rise to ferroresonance and negative damping, the full-wave rectified sine wave voltage rather stabilizes the circuit as no ferroresonance phenomenon appears.

Remarks About Flux-Current Graph

The graph showing the flux ϕ as a function of the current i is shown in Fig. 1. To facilitate comparison with other analytical results $\phi=\tanh i$. According to the proposed method, the shape of the graph can be taken directly from experimental data, similar to the B—H curve. For a nonlinear inductor, the $\phi-i$ graph may be also called a saturation curve, a term usually given to the B—H curve. The slope $\phi'=d\phi/di$ can be plotted from the $\phi-i$ graph. The reciprocal of ϕ' can also be obtained from the same $\phi-i$ graph. In fact, only $1/\phi'$ is needed in the solutions of the examples. In Fig. 1, both ϕ' and $1/\phi'$ are plotted with two different scales. For any value of i, $1/\phi'$ can be read from the reciprocal slope curve.

R, L, C Variable with Time

In the solutions of equations 17(A) and 18(A), one term of the acceleration function $G(t, i, q)$ involves the product of a current function $f(i)$ and a time function $F(t)$. Hence, the method given in this paper can be applied to any circuit with a nonlinear element $f(i)F(t)$. For a resistance varying with time, the term Ri in equation 5 should be generalized to $R(t)i$. Then $F(t)=R(t)$ and $f(i)=i$. Thus

$$\frac{di}{dt}=\frac{1}{\phi'(i)}[E-R(t)i-q/C]=G(t,i,q) \qquad (19)$$

$$\frac{di}{dq}=\frac{(1/\phi')[E-R(t)i-q/C]}{i}=\frac{G(t,i,q)}{i} \qquad (19A)$$

If $E=$constant and $(1/\phi')=\cosh^3 i/L$, then the most involved term in the G function is $(\cosh^2 i)iR(t)$ which is of the form $f(i)F(t)$ as $(\cosh^2 i)\sin t$ appeared in equation 18(A).

Let the capacitance element be represented by $S(t)$ or $1/C(t)$. Equation 5 will become

$$\phi'\frac{di}{dt}+R(t)i+S(t)q=E(t) \qquad (20)$$

or

$$\phi'\frac{di}{dt}+R(t)i+\frac{1}{C(t)}q=E(t) \qquad (21)$$

Then

$$\frac{di}{dt}=(1/\phi')[E-R(t)i-S(t)q]=G(t,i,q) \qquad (20A)$$

$$\frac{di}{dq}=\frac{(1/\phi')[E-R(t)i-S(t)q]}{i}=\frac{G(t,i,q)}{i} \qquad (20B)$$

or

$$\frac{di}{dt}=(1/\phi')[E-R(t)i-q/C(t)]=G(t,i,q) \qquad (21A)$$

$$\frac{di}{dq}=\frac{(1/\phi')[E-R(t)i-q/C(t)]}{i}=\frac{G(t,i,q)}{i} \qquad (21B)$$

Next let the flux ϕ be represented by $L(t)i$ instead of $\phi(i)$ in equation 4. Thus, instead of equation 5

$$L(t)\frac{di}{dt}+[L'(t)+R]\,i+\frac{q}{C}=E(t) \qquad (22)$$

where $L'(t)=dL/dt$. The acceleration function is now given by

$$\frac{di}{dt}=\left(\frac{1}{L(t)}\right)[E-R(t)i-q/C(t)]=G(t,i,q) \qquad (22A)$$

where $R(t)=L'(t)+R$, and $C(t)$ is taken from equation 21.

A Generalized Solution for Nonlinear R, L, and C

Let $\phi=\phi(i,t)=[L(i,\ t)i]$, R becomes $R(i, t)$, and q/C becomes $Q(q, t)$.

Then

$$\phi=\frac{d}{dt}[L(i,t)i]=\left[\frac{\partial L}{\partial i}i+L(i,t)\right]\frac{di}{dt}+\left[\frac{\partial}{\partial t}L(i,t)\right]i \qquad (23)$$

$$\phi'(i,t)\frac{di}{dt}+R'(i,t)i+Q(q,t)=E(t) \qquad (24)$$

where $\phi'(i,\ t)=(\partial L/\partial i)i+L(i,\ t)$, and $R'(i,t)=(\partial L/\partial t)+R(i,t)$.

The acceleration function is given by the following

$$\frac{di}{dt}=[1/\phi'(i,t)]\,[E(t)-R'(i,t)-Q(q,t)] \qquad (25)$$

The slope for the velocity plane plot based on the acceleration plane method is in general

$$\frac{di}{dq}=\frac{(1/\phi')[E(t)-R'(i,t)-Q(q,t)]}{i}=\frac{G(t,i,q)}{i} \qquad (26)$$

References

1. Nonlinear Analysis of Electro-Mechanical Problems, **Y. H. Ku.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 255, 1953, pp. 9–31.

2. Resonance Theory of Series Non-Linear Control Circuits, **E. G. Keller.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 225, 1938, pp. 561–77.

3. Analysis of Series Capacitor Application Problems, **J. W. Butler, C. Concordia.** *AIEE Transactions*, vol. 56, Aug. 1937, pp. 975–88.

4. The Reversion Method for Solving Nonlinear Differential Equations, **L. A. Pipes.** *Journal of Applied Physics*, New York, N. Y., vol. 23, 1952, pp. 202–07.

5. Acceleration Plane Method for Nonlinear Oscillations, **Y. H. Ku.** *Proceedings of the Symposium on Nonlinear Circuit Analysis* Polytechnic Institute of Brooklyn; New York, N. Y., vol. II, 1953, pp. 129–52.

6. A Mathematical Analysis of a Series Circuit Containing a Nonlinear Capacitor, **L. A. Pipes.** *AIEE Transactions*, vol. 72, pt. I, July 1953, pp. 238–44.

7. Discussion on Transient Analysis of A-C Machinery, **M. F. Gardner.** *AIEE Transactions*, vol. 48, 1929, pp. 714–15.

8. Transient Analysis of A-C Machinery—II, **Y. H. Ku.** *Journal*, Chinese Institute of Electrical Engineers, New York, N. Y., vol. 3, 1932, pp. 179–206. Also abstract in *Electrical Engineering*, vol. 51, 1932, p. 408.

9. Negative Damping of Electrical Machinery, **C. Concordia, G. K. Carter.** *AIEE Transactions* (*Electrical Engineering*), vol. 60, March 1941, pp. 116–19.

10. Self-Excitation of Induction Motors with Series Capacitors, **C. F. Wagner.** *AIEE Transactions*, vol. 60, 1941, pp. 1241–47.

11. Transient Torque-Angle Characteristics of Synchronous Machines, **W. V. Lyon, H. E. Edgerton.** *AIEE Transactions*, vol. 49, April 1930, pp. 686–98.

12. Unsymmetrical Self-Excited Oscillations in Certain Simple Non-Linear Systems, **J. G. Brainerd, C. N. Weygandt.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 24, 1936, pp. 914–22.

13. Notes on Modulation, **J. G. Brainerd.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 28, 1940, pp. 136–39.

14. Solutions of Mathieu's Equation, J. G.

Brainerd, C. N. **Weygandt.** *Philosophical Magazine*, London, England, vol. 30, 1940, pp. 458–77.

15. SINUSOIDAL VARIATION OF A PARAMETER IN A SIMPLE SERIES CIRCUIT, **F. J. Maginniss.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 29, 1941, pp. 25–28.

16. SINUSOIDAL VARIATION OF INDUCTANCE IN A LINEAR SERIES RLC CIRCUIT, **E. I. Hawthorne.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 39, 1951, pp. 79–81.

17. ON PARAMETRIC EXCITATION, **N. Minorsky.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 240, 1945, pp. 25–46. Also *Journal of Applied Physics*, New York, N. Y., vol. 22, 1951, p. 49.

18. ENERGY FLUCTUATIONS IN A VAN DER POL OSCILLATOR, **N. Minorsky.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 248, 1949, pp. 205–23.

19. STATIONARY SOLUTIONS OF CERTAIN NONLINEAR DIFFERENTIAL EQUATIONS, **N. Minorsky.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 254, 1952, pp. 21–42.

20. PERIODIC MOTIONS OF A NON-LINEAR DYNAMIC SYSTEM, **H. Serbin.** *Quarterly of Applied Mathematics*, Providence, R. I., vol. 8, 1950, pp. 296–303.

21. MATHEMATICAL ANALYSIS OF AN INDUCTIVELY LOADED PARALLEL-CONNECTED MAGNETIC AMPLIFIER, **L. A. Pipes.** *Journal of Applied Physics*, New York, N. Y., vol. 23, 1952, pp. 1358–61.

22. A METHOD FOR SOLVING THIRD AND HIGHER ORDER NONLINEAR DIFFERENTIAL EQUATIONS, **Y. H. Ku.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 256, 1953, pp. 229–44.

Analysis of Nonlinear Coupled Circuits

Y. H. KU
FELLOW AIEE

Synopsis: The method of high-order phase planes[1] or phase space is applied to the solution of nonlinear coupled circuits. Two simple examples are given to illustrate the application of this method. In the first example, a nonlinear resistance is introduced in the primary circuit. In the second example, a nonlinear inductance is used. It is possible to extend the application of the method to circuits with nonlinear capacitance and nonlinear mutual inductance. The method is applicable to any number of coupled circuits with nonlinear elements in any of the meshes or in the coupling between the meshes.

THERE has been considerable development in recent years in the analysis of nonlinear circuits and systems of a single degree of freedom or of a single loop. Also, many efforts have been made in solving second-order nonlinear differential equations by the phase-plane method. The present paper proposes to use the method of high-order phase planes[1] or the method of phase space for the solution of nonlinear coupled circuits or for nonlinear systems with more than one degree of freedom. The nonlinear differential equation thus encountered may easily reach third or higher orders. This method may also be considered as a means of solving nonlinear simultaneous equations. Though the method of phase space suggests a geometrical or topological approach, the actual analysis or solution is not limited to the graphical method. Thus high accuracy of the solution may be obtained with the help of computers.

General Outline of the Method

Given the differential equations of several coupled circuits with nonlinear elements, it is possible in many cases to eliminate all equations except one, which becomes a nonlinear differential equation of a higher order, higher than the order of any of the original set of equations.

Take a coupled circuit with two meshes or two loops, each with a resistance R and an inductance L and coupled by a mutual inductance M. While each loop by itself can be represented by a differential equation of the first order, the differential equation for the whole system is of the second order. Introduction of a nonlinear element in one of the loops will introduce nonlinearities in the differential equation, while the number of the degrees of freedom increases the order of the differential equation in the nonlinear case as in the linear case.

Consider the two mesh or two loop problem again with, say, an additional capacitance in the secondary circuit, see Fig. 1. The original set of differential equations are

$$R_1i_1+L_1pi_1-Mpi_2=e_1 \tag{1}$$

$$R_2i_2+L_2pi_2+q_2/C-Mpi_1=0 \tag{2}$$

where

i_1 and i_2 = the mesh or loop currents
$q_2=\int i_2dt$ = the charge in loop 2
e_1 = the applied voltage in loop 1
R_1 and R_2 = resistances
L_1 and L_2 = inductances
M = mutual inductance

Notice that any of the elements, R, L, or C, may be nonlinear. The symbol p stands for the time differential operator d/dt.

As these equations may involve nonlinear elements, the method of Laplace transforms or operational calculus cannot be used readily and the simultaneous solution of these equations may be made as follows:

1. If the nonlinear element is in equation 1, try to eliminate equation 2. If there are three or more simultaneous equations and the nonlinear element occurs in one of them, consider the equation with the nonlinear element as equation 1 and eliminate the rest.
2. After eliminating equation 2, differentiate both sides of equation 2 with respect to time t. This gives equation 2(A).
3. Assuming R_2, L_2, C, and M all linear constants, multiply all terms of equation 2(A) by Mp operationally. This is the same as differentiating both sides of equation 2(A) once again with respect to t and then multiply by a constant M. The result gives equation 2(B).
4. Multiply the different terms in equation 1 by the original operator before i_2 in equation 2(A). Since R_1 and L_1 may be nonlinear, both R_1i_1 and L_1pi_1 should be premultiplied by p or p^2 by taking the first or second derivative of R_1i_1 and L_1pi_1. This gives equation 1(A).
5. Since the operators before i_2 in both equations 1(A) and 2(B) are the same, eliminate these terms and combine the rest into equation 3, which is a nonlinear equation of a higher order than the order of either equations 1 or 2(A).
6. If there are three or more meshes or loops, first get the same operators for i_2, i_3, etc., and then eliminate all terms involving i_2, i_3, etc., to get a high-order nonlinear differential equation involving i_1 and its derivatives.
7. If there is nonlinearity in the coupling element, choose different mesh or loop currents so that one of the mesh or loop currents passes the nonlinear coupling element alone.

Similar procedures may be developed for the node-pair method as for the mesh method just outlined.

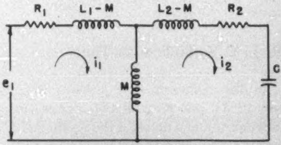


Fig. 1. Coupled circuit

Paper 54-515, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 14, 1954; made available for printing August 24, 1954.

Y. H. KU is with the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, Pa.

The equations mentioned in the foregoing are

$$(R_2p+L_2p^2+1/C)i_2-Mp^2i_1=0 \quad \text{(2A)}$$

$$(R_2p+L_2p^2+1/C)(Mp)i_2-M^2p^2i_1=0 \quad \text{(2B)}$$

$$(R_2p+L_2p^2+1/C)(R_1i_1+L_1pi_1)-(R_2p+L_2p^2+1/C)(Mp)i_2=(R_2p+L_2p^2+1/C)e_1 \quad \text{(1A)}$$

By adding equations 2(B) and 1(A)

$$(R_2p+L_2p^2+1/C)(R_1i_1+L_1pi_1)-M^2p^2i_1=f_1 \quad (3)$$

where $f_1=(R_2p+L_2p^2+1/C)e_1$.

If R_1 and L_1 are also linear constants, equation 3 can be easily solved by Laplace transforms as follows

$$(R_2s+L_2s^2+1/C)(R_1+L_1s)I_1(s)-M^2s^2I_1(s)=F_1(s)=(R_2s+L_2s^2+1/C)E_1(s) \quad (4)$$

where

s = complex variable
$I_1(s)$, $F_1(s)$, $E_1(s)$ = the Laplace transforms of $i_1(t)$, $f_1(t)$ and $e_1(t)$, and assuming no initial currents or charges

It is only necessary to find the inverse transform of $I_1(s)$ to get $i_1(t)$, for a given applied voltage $e_1(t)$.

Solution of a Third-Order Nonlinear Differential Equation

There are two general procedures for solving a high-order nonlinear differential equation like equation 3 where either R_1 or L_1 or both may be nonlinear. One procedure is to transfer all the nonlinear terms to the right-hand side of equation 3 resulting in

$$Z(p)i_1=f_1+n_1 \quad (5)$$

where

$Z(p)$ = linear operator
f_1 = as defined before
n_1 = an equivalent nonlinear forcing function involving i_1 and its derivatives

Solving $Z(p)i_1=f_1$ results in i_1 as a first approximation. After getting the first set of values of i_1, pi_1, p^2i_1, etc., the nonlinear function n_1 can be calculated. Then second and third approximations may be made for the solution of i_1.

The other procedure is to keep all the linear and nonlinear terms on the left side of equation 3 resulting in

$$Z(p)i_1+N_1(i_1)=f_1 \quad (6)$$

where $N_1(i_1)$ may involve a number of nonlinear terms such as i_1^2, $(pi_1)^2$, $(pi_1)^3$, $(pi_1)(p^2i_1)$, etc.

Then equation 3 or 6 can be solved by the method of high-order phase planes as outlined in the following.

Let $i_1=x_1$, $pi_1=v_1$, $p^2i_1=g_1$, and $p^3i_1=h_1$. Using the mechanical analogy, consider x_1 as displacement, v_1 as velocity, g_1 as acceleration, and h_1 as rate of acceleration. Equation 3 or 6 can then be written in the following form

$$K(v_1)h_1+F(v_1,x_1)g_1+F_1(v_1,x_1)v_1+F_2(x_1)=f_1(t) \quad (7)$$

where

$K(v_1)$ = a function of v_1
$F(v_1,x_1)$ and $F_1(v_1,x_1)$ = funtions of v_1 and x_1
$F_2(x_1)$ = a function of x_1

The symbol $f_1(t)$ is the same as f_1 defined before and depends upon $e_1(t)$ the applied voltage or forcing function.

Since $h_1=dg_1/dt$ and $g_1=dv_1/dt$, the following relation is obtained

$$\frac{dg_1}{dv_1}=\frac{h_1}{g_1} \quad (8)$$

This is a high-order phase plane equation determining the g_1-v_1 trajectory in the g_1-v_1 plane. It is similar to the first-order phase plane equation

$$\frac{dv_1}{dx_1}=\frac{g_1}{v_1} \quad (9)$$

used in the solution of second-order nonlinear differential equations.

In an unpublished 1953 AIEE Summer General Meeting paper the author presented a solution of the following equation representing a circuit with nonlinear inductance and capacitance

$$\phi(v_1)g_1+F(v_1,x_1)v_1+F_2(x_1)=f_1(t) \quad (10)$$

by the acceleration plane method.[2,3] Notice that in the right-hand member of equation 9, g_1 represents acceleration, while v_1 represents velocity. In other words, g_1 represents the second derivative of x_1, and v_1 represents the first derivative. The v_1-x_1 trajectory in the v_1-x_1 plane can be determined from the ratios of g_1 and v_1.

While in the solution of a second-order nonlinear differential equation one first-order phase plane equation is sufficient, in a third-order nonlinear differential equation both equations 8 and 9 are needed. In other words, there is a 3-dimensional phase space instead of a 2-dimensional phase plane. In this 3-dimensional phase space, three sets of trajectories can be traced: the $g-v$ trajectory in the $g-v$ plane, the $v-x$ trajectory in the $v-x$ plane, and the $g-x$ trajectory in the $g-x$ plane. However, two trajectories are sufficient, and all three trajectories are determined when any two of them are known. From equations 8 and 9

$$\frac{dg_1}{dx_1}=\frac{h_1}{v_1} \quad (11)$$

Example I

The method of solution is given by two examples. Example I is for a coupled circuit with nonlinear resistance in the first mesh or loop. Equations 1 and 2 become

$$R_1(i_1+i_1^2)+L_1pi_1-Mpi_2=e_1 \quad (12)$$

$$R_2i_2+L_2pi_2+q_2/C-Mpi_1=0 \quad (13)$$

Differentiating with respect to time t, equation 13 becomes

$$R_2pi_2+L_2p^2i_2+i_2/C-Mp^2i_1=0 \quad (14)$$

Eliminating i_2 from equations 12 and 14 the result is similar to equation 3

$$(R_2p+L_2p^2+1/C)(R_1+L_1p)i_1+(R_2p+L_2p^2+1/C)R_1i_1^2-M^2p^2i_1=f_1 \quad (15)$$

where $f_1=(R_2p+L_2p^2+1/C)e_1$. For $e_1=1$, $f_1=1/C$ equation 15 can be written as follows

$$(a_3p^3+a_2p^2+a_1p+a_0)i_1+b_2i_1(p^2i_1)+b_2(pi_1)^2+b_1i_1(pi_1)+b_0i_1^2=f_1 \quad (16)$$

where

$a_3=L_1L_2-M$
$a_2=R_1L_2+R_2L_1$
$a_1=R_1R_2+L_1/C$
$a_0=R_1/C$
$b_2=2R_1L_2$
$b_1=2R_1R_2$
$b_0=R_1/C$

From equation 12 on, R_1 is a constant, while the nonlinear resistance is given by $R_1(1+i_1)$.

Using the notations h_1, g_1, and v_1 to represent the third, second, and first derivatives of $x_1=i_1$, equation 16 becomes

$$a_3h_1+(a_2+b_2x_1)g_1+(a_1+b_1x_1)v_1+b_2v_1^2+b_0x_1^2+a_0x_1=f_1 \quad (17)$$

which gives

$$h_1=\frac{f_1-(a_2+b_2x_1)g_1-(a_1+b_1x_1)v_1-b_2v_1^2-b_0x_1^2-a_0x_1}{a_3} \quad (18)$$

Notice that equation 17 is a special form of equation 7 with $K(v_1)$ equal to a constant a_3. $F(v_1,x_1)$ is given by $(a_2+b_2x_1)$, $F_1(v_1,x_1)$ is given by $(a_1+b_1x_1+b_2v_1)$, and $F_2(x_1)$ is given by $(a_0+b_0x_1)x_1$.

In the actual solution of this example equation 8 and the reciprocal form of equation 9 are used such that two trajectories g_1-v_1 and x_1-v_1 can be plotted at the same time with v_1 as abscissa and g_1 and x_1 as ordinates.

The phase plane equations are

$$\frac{dg_1}{dv_1}=\frac{h_1}{g_1}=\frac{f_1-(a_2+b_2x_1)g_1-(a_1+b_1x_1)v_1-b_2v_1^2-b_0x_1^2-a_0x_1}{a_3g_1} \quad (19)$$

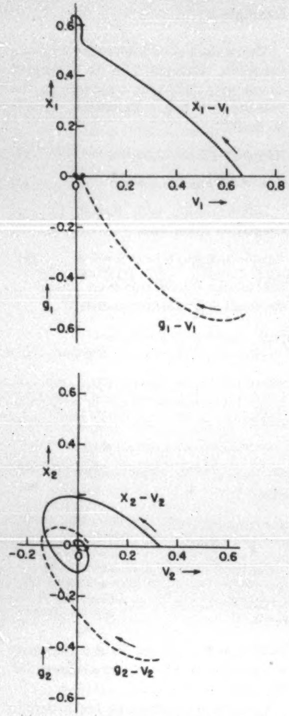


Fig. 2. Trajectories for example I

$$\frac{dx_1}{dv_1}=\frac{v_1}{g_1} \tag{20}$$

The next step is to get all the initial values. With values of h_1, g_1, and v_1 known, the initial slopes of the $g-v$ and $x-v$ trajectories can be calculated by the relations given in equations 19 and 20. Suppose these slopes in the $g-v$ plane and in the $x-v$ plane are drawn from the initial points (g_1,v_1) and (x_1,v_1). Taking the velocity (or first derivative) v as the horizontal axis, a small step Δv can be assumed. From the known initial slopes

$$\begin{aligned}\Delta g_1&=(h_1/g_1)_{t=0}\Delta v_1\\ \Delta x_1&=(v_1/g_1)_{t=0}\Delta v_1\end{aligned} \tag{21}$$

Substituting the new set of values of v_1, g_1, and x_1 into equation 18, the corresponding value of h_1 is obtained at the end of the interval corresponding to Δv. The new set of values of h_1, g_1, v_1, and x_1 will give new slopes. Here time is implicit. Instead of working with Δt, Δv is used in this example. In working with Δt, a value of h_1 has to be assumed at the end of the interval and an average value of h has to be obtained during the interval. Then $h_{av}(\Delta t)=\Delta g$. Taking the average value of g results in $g_{av}(\Delta t)=\Delta v$. Taking the average value of v, $v_{av}(\Delta t)=\Delta x$. However, equation 17 or 18 must be satisfied. By substituting in all the values of g, v, and x thus found into equation 18, a calculated value of h is obtained which must be compared to the assumed value used at the start. Suitable corrections are then made and a new set of calculations must be made. The advantage of using Δv instead of Δt is that, besides the initial values, the additional knowledge about the initial slopes of the trajectories is also utilized. Many ways of checking the intermediate slopes during the interval can be developed. The interested reader may try out his check by working on a first-order phase plane equation as given in equation 9. This problem was assigned to the graduate students in a class on nonlinear analysis, and satisfactory ways of determining the slope within a given interval have been demonstrated by the students themselves.

To facilitate further discussion, the constants will be given simple numerical values. Let $R_1=R_2=1$, $L_1=L_2=2$, $M=1$, and $C=1$. Assume also $e_1=1$, hence, f_1 is also equal to 1. Equation 18 then becomes

$$h_1=\frac{\begin{array}{r}1-4(1+x_1)g_1-(3+2x_1)v_1-\\ 4v_1^2-(x_1+x_1^2)\end{array}}{3} \tag{22}$$

The initial values are

$$\begin{aligned}&x_1=0,\ v_1=2/3,\ g_1=-5/9,\ h_1=-5/27\\ &x_2=0,\ v_2=1/3,\ g_2=-4/9,\ h_2=-1/27,\ q_2=0\end{aligned}$$

The initial slopes are

$$\frac{dg_1}{dv_1}=\frac{h_1}{g_1}=1/3,\ \frac{dx_1}{dv_1}=\frac{v_1}{g_1}=-6/5$$

$$\frac{dg_2}{dv_2}=\frac{h_2}{g_2}=1/12,\ \frac{dx_2}{dv_2}=\frac{v_2}{g_2}=-3/4$$

As $\Delta t=\Delta x/v_{av}$ can be obtained from the $v-x$ trajectory, $v(t)$ and $x(t)$ can be plotted. Knowing $v(t)$ and $x(t)$, at any instant, the corresponding values of $g(t)$ and $h(t)$ can be obtained and then plotted. The g_1-v_1, x_1-v_1, g_2-v_2, and x_2-v_2 trajectories for this numerical example are shown in Fig. 2. The h_1-x_1, g_1-x_1 and x_2-q_2 trajectories are shown in Fig. 3.

Marvin Gottlieb[4] used both sets of phase plane equations: the first set as given by equations 19 and 20; the second set as given by equation 9 and

$$\frac{dv_2}{dx_2}=\frac{g_2}{v_2} \tag{23}$$

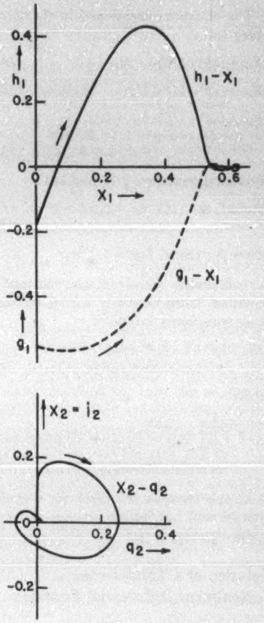


Fig. 3. Trajectories for example I

Just to complete the example, the following relations between the primary and secondary quantities are

$$h_2=2h_1+(1+2x_1)g_1+2v_1^2 \tag{24}$$

$$g_2=2g_1+v_1(1+2x_1) \tag{25}$$

$$v_2=2v_1+x_1(1+x_1)-1 \tag{26}$$

$$x_2=-3g_1-(4v_1+x_1)(1+x_1)+1 \tag{27}$$

$$q_2=3g_1+v_1(1+4x_1)-x_1(1+x_1) \tag{28}$$

It is not necessary to express the secondary quantities in terms of primary quantities only. After certain secondary quantities such as v_2 and x_2 are found, h_2 and g_2 may be obtained in terms of some primary quantities and the known values of, say, v_2 and x_2.

Example II

In example II, a nonlinear inductance in the first mesh or loop is assumed. Equation 1 becomes

$$R_1i_1+\frac{d\phi_1}{dt}+Mpi_1-Mpi_2=e_1 \tag{29}$$

Let

$$\phi_1=L_1\tanh i_1=L_1\alpha$$

Then $p\phi_1=d\phi_1/dt=(L_1/\cosh^2 i_1)\ (pi_1)$

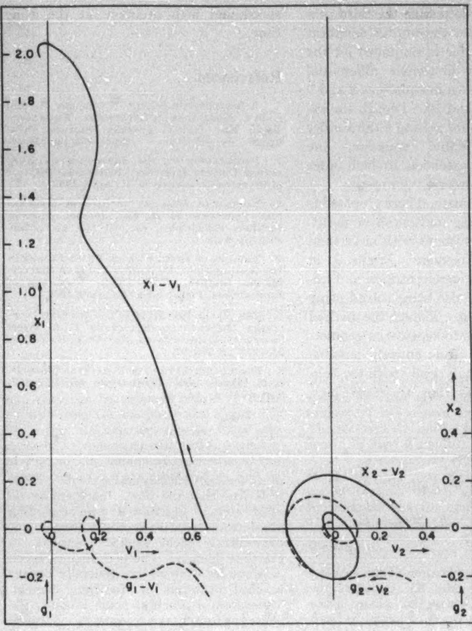


Fig. 4 (left). Trajectories for example II

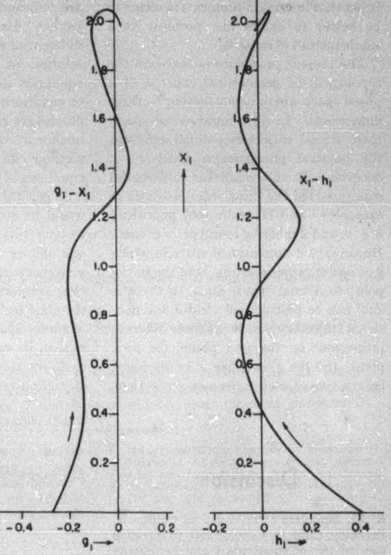


Fig. 5 (above). Trajectories for example II

$=L_1\beta\,(pi_1)=L_1\beta v_1$. Differentiating once, equation 29 becomes

$$R_1(pi_1)+p(L_1\beta v_1)+Mp^2i_1-Mp^2i_2=\dot{e}_1 \qquad (30)$$

where $\dot{e}_1=de_1/dt=pe_1$. Now $p(L_1\beta v_1)=L_1\beta(g_1-2\alpha v_1^2)$ as $p\beta=-2\alpha\beta v_1$. Equation 30 can be written as

$$R_1v_1+L_1\beta(g_1-2\alpha v_1^2)+Mg_1-Mg_2=\dot{e}_1 \qquad (31)$$

Differentiating once again, equation 31 becomes

$$R_1g_1+L_1\beta(h_1-4\alpha v_1g_1)+L_1(g_1-2\alpha v_1^2)(-2\alpha\beta v_1)-2L_1\beta^2v_1^3+Mh_1-Mh_2=\ddot{e}_1 \qquad (32)$$

where $\ddot{e}_1=p^2e_1$. Equation 32 can be simplified to

$$(L_1\beta+M)h_1+(R_1-6\alpha\beta L_1v_1)g_1+2\beta L_1(2\alpha^2-\beta)v_1^3-Mh_2=\ddot{e}_1 \qquad (33)$$

In this example, the equation for the second mesh or loop is the same as equation 2 with constant coefficients. Applying the operator $(R_2p+L_2p^2+1/C)$ to equation 29 and eliminating the i_2 term as in equation 3

$$L_2(L_1\beta+M)h_1+(R_1L_2+R_2L_1\beta-6\alpha\beta L_1L_2v_1)g_1+2\beta L_1L_2(2\alpha^2-\beta)v_1^3-M^2h_1+(MR_2)g_1-2\alpha\beta L_1R_2v_1^2+(R_1R_2+M/C+L_1\beta/C)v_1+(R_1x_1)/C=(R_2\dot{e}_1+L_2\ddot{e}_1+e_1/C)=f_1 \qquad (34)$$

Collecting terms results in

$$(a_3+b_3\beta)h_1+(a_2+b_2\beta-6\alpha\beta b_3v_1)g_1+(a_1+b_1\beta)v_1-2\alpha\beta b_2v_1^2+2b_3\beta(2\alpha^2-\beta)v_1^3+a_0x_1=f_1 \qquad (35)$$

where

$a_3=(L_2-M)M$ $\quad$ $a_1=R_1R_2+M/C$
$b_3=L_1L_2$ $\quad$ $b_1=L_1/C$
$a_2=R_1L_2+R_2M$ $\quad$ $a_0=R_1/C$
$b_2=R_2L_1$

The following numerical values are introduced: $R_1=R_2=0.5$, $L_1=M=1$, $L_2=2$, $C=1$, and $e_1=1$. Then, from equation 35

$$h_1=\frac{1-(1.5+0.5\beta-12\alpha\beta v_1)g_1-(1.25+\beta)v_1+\alpha\beta v_1^2+4\beta(2\alpha^2-\beta)v_1^3-0.5x_1}{1+2\beta} \qquad (36)$$

The initial values are

$x_1=0,\ x_2=0,\ q_2=0$

$v_1=2/3,\ v_2=1/3$

$g_1=-5/18,\ g_2=-2/9$

$h_1=67/162,\ h_2=31/324$

The initial slopes are

$$\frac{dg_1}{dv_1}=\frac{h_1}{g_1}=-67/45 \qquad \frac{dg_2}{dv_2}=\frac{h_2}{g_2}=-31/72$$

$$\frac{dx_1}{dv_1}=\frac{v_1}{g_1}=-12/5 \qquad \frac{dx_2}{dv_2}=\frac{v_2}{g_2}=-3/2$$

The relations between the secondary and primary quantities are

$$h_2=h_1/2-g_2/4-v_2/2 \qquad (37)$$

$$g_2=v_1/2+(1+\beta)g_1-2\alpha\beta v_1^2 \qquad (38)$$

$$v_2=x_1/2+(1+\beta)v_1-1 \qquad (39)$$

$$x_2=g_1-2g_2-v_2/2 \qquad (40)$$

$$q_2=v_1-2v_2-x_2/2 \qquad (41)$$

Notice that with the primary quantities known, g_2 and v_2 can be obtained first. The x_1-v_1, g_1-v_1, x_2-v_2, and g_2-v_2 trajectories are shown in Fig. 4, while the x_1-g_1 and x_1-h_1 trajectories are shown in Fig. 5.

Conclusion

The method of high-order phase plane or phase space is applied to the solution of nonlinear coupled circuits. There are two general procedures for attacking nonlinear differential equations of higher order than the second. While the procedure of transferring all nonlinear terms to the forcing-function side as additional forcing functions can utilize known methods of Laplace transforms and operational calculus, the procedure here presented for the first time is based on the belief that in certain problems it seems to be better to attack the problem as a whole instead of in parts.

The present procedure is based on the topological or geometrical concept of a phase space which is not limited to three dimensions. In an n-dimensional space, there are n-1 trajectories which represent the physical phenomenon which is of concern. If the nonlinear differential equation is of the third order, as given in examples I and II, really four quantities h, g, v, and x must be found or evaluated. However, a 3-dimensional space in which g, v, and x values starting from the initial point to a final steady-state (if there is one) can be plotted will yield a 3-dimensional trajectory which has three different projections on the $g-v$ plane, the $v-x$ plane, and the $g-x$ plane. As discussed before, two of such trajectory projections are sufficient to determine the third projection. Hence, a third-order nonlinear differential equation is simplified for the solution of two first-order differential equations, as shown in equations 8 and 9 or equations 19 and 20. Though the examples are given for getting a third-order nonlinear differential equation, the method can be extended to high-order equations of the order 4, 5, or more.

In general, the method here proposed is useful in attacking multimesh or multiloop circuits, or systems with more than one degree of freedom. Problems in servomechanisms, aerodynamics, and nuclear reactors are also being solved using the same technique. Though the method of phase space is topological or geometrical in its origin, it is entirely possible to do the steps analytically with the help of computers and thus to insure high speed and high accuracy at the same time.

References

1. A METHOD FOR SOLVING THIRD AND HIGHER ORDER NONLINEAR DIFFERENTIAL EQUATIONS, **Y. H. Ku.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 256, no. 3, Sept. 1953, pp. 229–44.

2. PROCEEDINGS OF THE SYMPOSIUM ON NONLINEAR CIRCUIT ANALYSIS. Polytechnic Institute of Brooklyn, New York, N. Y., April, 1953.

3. NONLINEAR ANALYSIS OF ELECTRO-MECHANICAL PROBLEMS, **Y. H. Ku.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 255, no. 1, Jan. 1953, pp. 9–31.

4. THE USE OF HIGHER ORDER PHASE PLANES IN THE ANALYSIS OF A NONLINEAR COUPLED CIRCUIT, **Marvin Gottlieb.** *M.S. Thesis*, University of Pennsylvania, Philadelphia, Pa., June 1953.

5. THE REVERSION METHOD FOR SOLVING NONLINEAR DIFFERENTIAL EQUATIONS, **L. A. Pipes.** *Journal of Applied Physics*, New York, N. Y., vol. 23, 1952, pp. 202–07.

6. FORCED OSCILLATIONS OF NONLINEAR CIRCUITS, **L. A. Pipes.** *AIEE Transactions*, vol. 73, pt. I, Sept. 1954, pp. 351–58.

◆

Discussion

R. M. Saunders (University of California, Berkeley, Calif.): I should like to question the validity of equations 1 and 2. Should not the statement be expanded to reflect the fact that L_1, M, L_2, and C are functions of current for the example used? Thus, equations 1 and 2 should read

$$R_1i_1+L_1pi_1-Mpi_2+i_1pl_1-i_2pM=e_1$$

$$R_2i_2+L_2pi_2+\int\frac{1}{C}i_2dt-Mpi_1+i_2pl_2-i_1pM=0$$

Regarding the application of the author's results, I assume Mr. Ku considers the problem solved when the phase space portrait is obtained. Not so with those of us dealing with physical systems. While it is true that the phase space curves do yield certain qualitative data regarding absolute and relative stability, they do not conveniently yield explicit solutions in the time domain. The performance specifications for systems are generally always couched in terms of the time domain. Thus, from a practical point of view, the phase portrait is not as easy an approach as some other technique. In my own limited experience, the step-by-step solution in the time domain is as easy as the method presented by Professor Ku.

There is an ever-growing feeling among those active in the nonlinear field that they must resign themselves to the use of computers, either analogue or digital, to effect solutions for nonlinear differential equations. When one becomes involved in a computer study, he will stick to the time domain.

Y. H. Ku: I appreciate the discussion of Mr. Saunders. Regarding equations 1 and 2 in the paper, Mr. Saunders' suggested forms are more general and should be used. For the examples given in the paper, the form of equation 1 can also be used if R and L are considered nonlinear. For example I, equation 12 gives a nonlinear resistance $R(i_1)=R_1(1+i_1)$. For example II, by substituting $L_1 \tanh i_1$ for ϕ_1 in equation 29, one obtains $d\phi_1/dt=(L_1/\cosh^2 i_1)(pi_1)$. Thus, nonlinear inductance $(M+L_1\beta)$ is obtained, where β stands for $1/\cosh^2 i_1$.

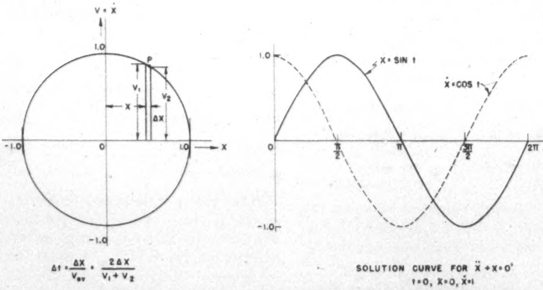


Fig. 6. Relation between phase plane plot and solution curves x(t) and v(t)

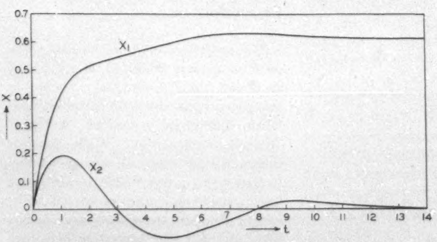


Fig. 7. Solution curves $x_1(t)$ and $x_2(t)$ for example I

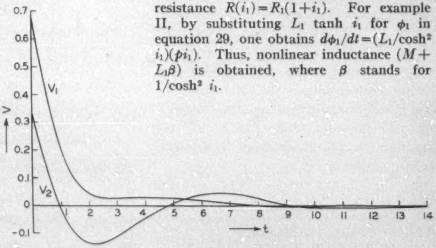
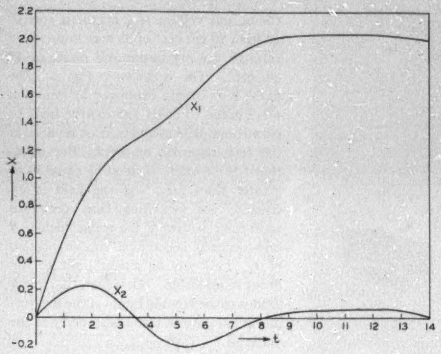


Fig. 8. Solution curves $v_1(t)$ and $v_2(t)$ for example I



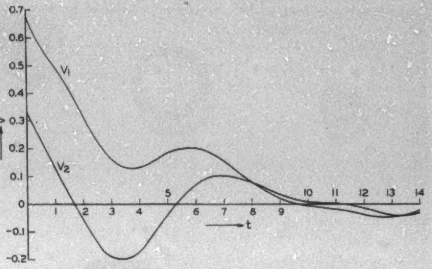


Fig. 9 (left). Solution curves $x_1(t)$ and $x_2(t)$ for example II

Fig. 10 (above). Solution curves $v_1(t)$ and $v_2(t)$ for example II

Regarding the solution curves $x(t)$ and $v(t)$, they can be obtained from any given v-x trajectory (or the phase space portrait) by noting the relation $\Delta t = \Delta x / v_{av}$, where v_{av} is the average value of v_1 and v_2, as shown in Fig. 6. In numerical calculations, the increments of Δx are tabulated as well as the values of the velocities (or first derivatives). Thus, Δt can be easily calculated and also the corresponding values of x, v, and t. The $x(t)$ and $v(t)$ curves for example I are shown in Figs. 7 and 8. The $x(t)$ and $v(t)$ curves for example II are shown in Figs. 9 and 10.

In a recent paper[1] credit was given to me for my suggestion of the means of carrying out the steps of the "graphical analysis" by programming the analysis on a digital computer. I believe that the graphical analysis presented in this paper can also be suitably programmed on computers.

REFERENCE

1. TRAVELING WAVE PROTECTION PROBLEMS—1. A GRAPHICAL ANALYSIS WITH SPECIAL APPLICATION TO THE SWITCHING OF LINES AND CABLES, E. W. Boehne. *AIEE Transactions*, vol. 73, pt. III, Aug. 1954, pp. 920–28.

A Comparison of Metals and Ferrites for High-Speed Pulse Operation

D. R. BROWN
NONMEMBER AIEE

D. A. BUCK
NONMEMBER AIEE

N. MENYUK
NONMEMBER AIEE

Synopsis: Metals and ferrites with rectangular hysteresis loops are used in magnetic amplifiers, switching networks, and data-storage systems. In these applications, the large-signal pulse response is most important. To reduce the effect of eddy currents, metals have been rolled to ultrathin gauge and high-resistivity ferrites have been used. By these techniques, eddy currents have been reduced to the point where relaxation effects limit the response time in both cases. The relative merits of metals and ferrites indicate that at present ultrathin metal-ribbon cores are better for stepping registers and switching circuits, while ferrite cores are better for use in coincident-current memories.

BECAUSE of their high reliability, magnetic cores with rectangular hysteresis loops are being used in high-speed memory and switching circuits. In these applications, magnetic fields which exceed the coercive force of the material are applied as step functions (rise times of approximately 0.2 microsecond). Since both ferrites and ultrathin metal-ribbon cores are frequently used, a comparison of their properties is of interest. This comparison is based on magnetic materials which are currently available and have been investigated in this laboratory.

Physical Description and Fabrication Technique

Rectangular-loop metal cores are prepared by a process of cold-rolling and annealing. Eddy currents are reduced by rolling the magnetic metal to an ultrathin ribbon from which ring-shaped cores are made by wrapping a strip of the ribbon around a bobbin. Since the material is strain-sensitive, this wrapping usually precedes the anneal. Metal cores, therefore, are wrapped on ceramic bobbins which are able to withstand the annealing temperature; see Fig. 1. A typical core might have 40 wraps of 1/8-mil (0.000125 inch) molybdenum-Permalloy ribbon on a 1/8-inch-wide and 3/16-inch-diameter bobbin.

Ferrites are magnetic ceramics which are prepared by pressing a powder into the desired shape and sintering it at high temperature. Since magnesium-manganese ferrites have d-c resistivities of the order of 10^{12} that of metals, eddy-current effects are negligible and laminations are not required. Different size ferrite cores prepared in this way are also shown in Fig. 1. The smallest core shown, with an outside diameter 0.090 inch and inside diameter of 0.050 inch, is used in the coincident-current memory of Whirlwind I at the Massachusetts Institute of Technology.

Uniformity is a major problem in both metal-ribbon and ferrite cores. The poor uniformity of metal-ribbon cores has limited their application. However, the uniformity of ferrite cores has improved remarkably since 1952, and this improvement is continuing. Ferrite cores can be produced on a large scale at a lower cost than ultrathin metal-ribbon cores. For

---

Paper 54-361, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted July 29, 1954; made available for printing August 6, 1954.

D. R. BROWN, D. A. BUCK, and N. MENYUK are with the Massachusetts Institute of Technology, Cambridge, Mass.

The research in this document was supported jointly by the Army, Navy, and Air Force under contract with the Massachusetts Institute of Technology.

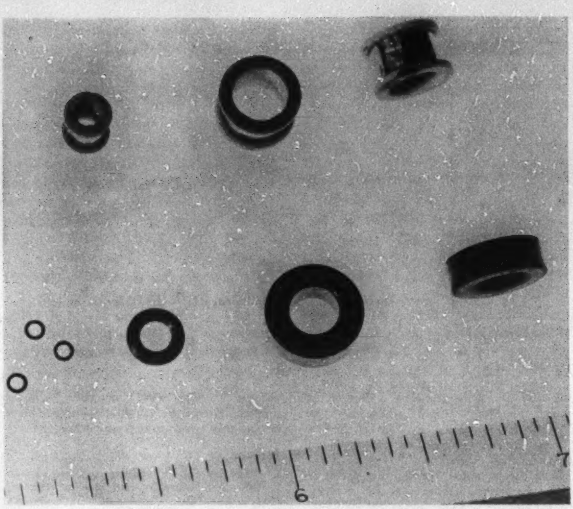


Fig. 1. Typical ferrite and ultrathin metal-ribbon cores

equipment requiring a large number of cores, this can be a decisive factor.

Energy-Loss Factors

The three major energy losses upon reversing the magnetization of a ferromagnetic core are hysteresis loss, eddy-current loss, and relaxation loss. The first two losses are familiar; the relaxation loss is essentially a damping loss manifested by the delay of the electron spin vectors in aligning themselves in the direction of the applied field. This phenomenon has been discussed extensively.[1-3]

EDDY CURRENTS

The reduction of eddy-current loss through the use of thinner metal ribbons permitted higher repetition rates until ultrathin ribbons below 1 mil in thickness were obtained. Further reduction to 1/4-mil and 1/8-mil ribbons did not permit the higher repetition rate expected. This indicates that losses other than eddy-current losses are predominant. In ferrites, high resistivities make eddy-current losses negligible.

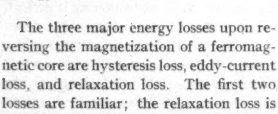


Fig. 2. Definition of switching time

HYSTERESIS

The hysteresis loss per unit volume per cycle of rectangular ferrites is approximately equal to $4I_sH_c$, where I_s is the saturation magnetization and H_c is the coercive force. In this case $I_s = B_s/4\pi$, where B_s is the saturation flux density in gausses, or electromagnetic units (emu). In magnesium-manganese ferrites, I_s is approximately 10^2 (emu) and H_c is approximately 1 oersted (emu); in molybdenum-Permalloy cores, I_s is approximately 10^3 (emu) and H_c is approximately 0.1 oersted (emu). Therefore, the hysteresis loss is approximately 400 ergs per centimeter^3 per cycle in both cases.

The Switching Coefficient

DEFINITION

The switching time τ of rectangular-loop materials is observed to be inversely proportional to the applied field, according to the relationship

$$(H-H_0)\tau = S_w \tag{1}$$

where H is the applied field, H_0 is the threshold field for irreversible domain-wall motion, and S_w is a constant defined as the switching coefficient. The switching time τ is defined as the time required for the output voltage of a magnetic core to go from 10 per cent of its maximum value through the maximum and down to 10 per cent. This is shown in Fig. 2. The linear relationship expressed by equation 1 is limited to fields sufficiently high to permit complete magnetization reversal of the ferromagnetic material. For equation 1 to be valid, H must be equal to or greater than $2H_c$. As the field is increased, the switching time decreases until the rise time of the applied field (0.2 microsecond) becomes a sizable portion of τ. This sets an upper limit to the linear relationship. However, this upper limit is caused by the limits of the measuring technique and is not a property of the material.

Equation 1 can be expressed in the form

$$H = \frac{S_w}{\tau} + H_0 \tag{2}$$

which is the equation of a straight line of slope S_w and intercept H_0.

EXPERIMENTAL RESULTS

Curves of the inverse switching time as a function of the applied field are shown for 1/8-mil and 1/4-mil 4-79 molybdenum-Permalloy cores and for a magnesium-manganese-ferrite core (Ferramic *MF-1312B*) in Fig. 3; the important parameters are listed in Table I. Note that the value of S_w is of the same order of magnitude for materials as different as molybdenum-Permalloy and magnesium-manganese ferrite. This is even more striking in view of additional experiments conducted in this laboratory with other ferrites, including magnesium-manganese ferrites of various compositions and several nickel-zinc ferrites. In every case, the resultant S_w was found to lie within a factor of 2 of the value given in the table for Ferramic *MF-1312B*. The combination of a lower threshold field and lower switching coefficient permits the 4-79 molybdenum-Permalloy to switch

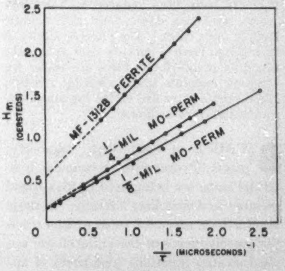


Fig. 3. Switching characteristics of metals and ferrites

approximately twice as fast as the ferrite at a given applied field value.

THEORETICAL CONSIDERATIONS

The experimental value of the switching coefficient is the sum of two effects, as may be expressed by the equation

$$S_w = S_w^{(r)} + S_w^{(e)} \tag{3}$$

where $S_w^{(r)}$ is the contribution to the switching coefficient caused by relaxation effects, and $S_w^{(e)}$ is the eddy-current contribution. These contributions to the switching coefficient can be related to the physical constants of the material by the equations[4]

$$S_w^{(r)} = \frac{I_s{}^{\Lambda d}}{(\gamma^2 I_s{}^2 + \Lambda^2)<\cos\theta>}\sqrt{\frac{K}{A}} \approx \frac{\Lambda d}{\gamma^2 I_s <\cos\theta>}\sqrt{\frac{K}{A}} \tag{4}$$

$$S_w^{(e)} = \frac{8\pi^2 I_s r_m{}^2}{\rho c^2 <\cos\theta>^3} \tag{5}$$

In equation 4, Λ represents the relaxation frequency of the material, I_s the saturation magnetization, d the maximum distance a domain wall moves during a magnetization reversal, K the effective anisotropy constant, γ the magneto-mechanical ratio, $<\cos\theta>$ the mean value of the cosine of the angle between the applied field and the direction of easy magnetization, and A is the exchange parameter. The parameters Λ, K, A, and d are unknown for the materials in question. However, a reasonable estimate can be made of Λ and K by comparison with the known values of these parameters in similar materials. The distance d can be estimated from the grain structure of the material and A can be approximated by $A \sim kT_c/a$ where T_c is the Curie temperature; k is the Boltzmann constant; and a is the lattice parameter. Calculations of $S_w^{(r)}$ based on these approximations are in agreement with experimental results.

In equation 5, valid only for thickness less than 1 mil, r_m is one-half the ribbon thickness, ρ is the resistivity, and c is the velocity of light. For 1/8-mil and 1/4-mil molybdenum-Permalloy ribbon, $I_s = 700$ (emu), $\rho = 6\times10^{-17}$ statohm-centimeter, $<\cos\theta> \approx 1$ (since this material is grain oriented), and $r_m = 1.6\times10^{-4}$ centimeter for 1/8-mil and 3.2×10^{-4} centimeter for the 1/4-mil ribbon. Substitution of these values in equation 5 yields

$$\begin{aligned} S_w{}^e({}^1/_8\text{-mil}) &= 0.3\times10^{-7} \text{ oersted-second} \\ S_w{}^e({}^1/_4\text{-mil}) &= 1.2\times10^{-7} \text{ oersted-second} \end{aligned} \tag{6}$$

Table I. Switching Characteristics of Metals and Ferrites

Material	S_w, Oersted (Emu)-Second	H_0, Oersted (Emu)
1/8-mil 4–79 molybdenum-Permalloy	0.55×10^{-6}	0.14
1/4-mil 4–79 molybdenum-Permalloy	0.63×10^{-6}	0.14
Ferramic *MF-1312B*	1.02×10^{-6}	0.52

Since $S_w{}^r$ is independent of the thickness of the material, the variation of S_w between 1/8-mil and 1/4-mil molybdenum-Permalloy, as given in Table I, is entirely caused by the change in the eddy-current contribution. $S_w{}^e$ varies as the square of the thickness, so

$$S_w{}^e({}^1/_4\text{-mil}) = 4S_w{}^e({}^1/_8\text{-mil})$$

and

$$S_w({}^1/_4\text{-mil}) = S_w{}^r + 4S_w{}^e({}^1/_8\text{-mil}) = 6.3\times10^{-7} \text{ oersted-second}$$

$$S_w({}^1/_8\text{-mil}) = S_w{}^r + S_w{}^e({}^1/_8\text{-mil}) = 5.5\times10^{-7} \text{ oersted-second}$$

Therefore

$$\begin{aligned} S_w{}^e({}^1/_8\text{-mil}) & 0.27\times10^{-7} \text{ oersted-second} \\ S_w{}^e({}^1/_4\text{-mil}) & 1.07\times10^{-7} \text{ oersted-second} \end{aligned} \tag{7}$$

The theoretical and experimental results in equations 6 and 7 are in reasonable agreement. They show that for 1/8-mil ribbon the eddy-current loss is reduced to approximately 5 per cent of the relaxation loss.

ENERGY LOSS

The energy loss per cubic centimeter per cycle is related to the switching coefficient by the equations[4]

$$\text{eddy-current loss} = \frac{2I_s <\cos\theta>}{\tau} S_w{}^e \tag{8}$$

$$\text{relaxation loss} = \frac{2I_s}{\tau} S_w{}^r \tag{9}$$

Since I_s is approximately 700 (emu) for 4-79 molybdenum-Permalloy and approximately 200 (emu) for magnesium-manganese ferrite, the total relaxation and eddy-current loss for these materials when switching in 1 microsecond is approximately 770 ergs per cubic centimeter per cycle for the metal, and 400 ergs per cubic centimeter per cycle for the ferrite.

Heating Effects

For the applications discussed here, power dissipation is directly proportional to the repetition frequency as long as the peak amplitude of the applied field is unchanged. However, a complete magnetization reversal per half-cycle requires that the switching time be less than a half period. At higher frequencies, this necessitates higher applied fields and increased power loss. Although the power dissipation per cubic centimeter per cycle of the metals and ferrites is comparable at room temperature, operation at high frequencies gives rise to heating effects which limit the operation of the magnetic core as a switching device. Therefore, the thermal behavior of ferrites and metals must be compared under these conditions.

OUTPUT SIGNAL

The higher flux density of the metal ribbons permits the use of a smaller cross section, hence smaller volume, of metal than ferrite to achieve the same flux change. Thus, the power dissipation per core is somewhat smaller in the metals. Because of the lower switching coefficient, the metals require less magnetomotive force to obtain the same time-rate-of-change of flux.

THERMAL CONDUCTIVITY

The thermal conductivity of molybdenum-Permalloy is high. Therefore, any cooling device applied to the surface of the core will be effective in cooling the entire volume. The thermal conductivity of ferrites is low, so surface cooling will only succeed in setting up a thermal gradient within the sample.

TEMPERATURE SENSITIVITY

Since molybdenum-Permalloy has a higher Curie temperature (460 degrees centigrade) than the ferrites (300 degrees centigrade), the metal core is less sensitive to small temperature changes in the vicinity of room temperature. For a specific multiposition magnetic-core switch a temperature rise of 90 degrees centigrade was found permissible for 4-79 molybdenum-Permalloy as compared to a 20-degree-centigrade rise for Ferramic *MF-1312B*.

Conclusions

The factors noted previously favor the use of ultrathin metal-ribbon cores for stepping registers and switching circuits. Stepping registers using these ribbons have been operated up to 500,000 cycles per second. For high-speed coincident-current memory applications of the type introduced by J. W. Forrester,[5] ferrite cores are better. In this application, the applied field is limited to twice the coercive force. Because of the low coercive force of molybdenum-Permalloy cores, their use in this application is limited to the region of slow switching (switching times of from 10 to 20 microseconds). The higher coercive force of the ferrites permits switching times of the order of 1 microsecond. Also, a large coincident-current memory requires tens of thousands of uniform cores. The poor uniformity of metal cores makes the selection of this many metal cores difficult.

Future developments may alter the foregoing conclusions. Ultrathin metals with a coercive force comparable to that of the ferrites have been reported, and improved fabrication techniques may overcome the uniformity problem. On the other hand, ferrites with a coercive force of 0.2 oersted have been produced, and progress is being made in obtaining ferrites of higher flux density than that of the magnesium-manganese ferrites discussed in the foregoing. However, the ferrimagnetic nature of ferrites limits the improvement that can be made in this direction. Developmental work involving the use of deposited and evaporated films is now going on in several laboratories. This may lead to improvements in the near future.

References

1. ON THE THEORY OF THE DISPERSION OF MAGNETIC PERMEABILITY IN FERROMAGNETIC BODIES, **L. Landau, E. Lifshitz.** *Physikalische Zeitschrift der Sowjetunion*, Kharkov, USSR, vol. 8, 1935, pp. 153–169.

2. PHYSICAL THEORY OF FERROMAGNETIC DOMAINS, **C. Kittel.** *Reviews of Modern Physics*, New York, N. Y., vol. 21, 1949, pp. 541–583.

3. MOTION OF A FERROMAGNETIC DOMAIN WALL IN FE_3O_4, **J. K. Galt.** *Physical Review*, New York, N. Y., vol. 85, 1952, pp. 664–669.

4. A THEORY OF DOMAIN CREATION, COERCIVE FORCE, AND FLUX REVERSAL IN POLYCRYSTALLINE FERROMAGNETICS, **J. B. Goodenough, N. Menyuk.** *Lincoln Laboratory Technical Report No. 40*, Massachusetts Institute of Technology, Cambridge, Mass., Jan. 1954. Also to be published in *Journal of Applied Physics*, New York, N.Y.

5. DIGITAL INFORMATION STORAGE IN THREE DIMENSIONS USING MAGNETIC CORES, **J. W. Forrester.** *Journal of Applied Physics*, New York, N. Y., vol. 22, 1951, pp. 44–48.

◆

Discussion

H. Ekstein and **T. L. Gilbert** (Armour Research Foundation of Illinois Institute of Technology, Chicago, Ill.): This discussion concerns only the theoretical considerations in so far as equations 4 and 5 and the model from which they are derived, are applied to ferromagnetic alloy tapes.

The authors have based their theory on a model wherein remagnetization starts by the formation of spikes of reversed magnetization at grain boundaries throughout the volume of the material. These spikelike reversal domains then grow, until flux reversal is complete. The model was obtained by deductive reasoning, which may well be sound for quasi-static processes in polycrystalline ferromagnetic alloys or for ferrites.[4] It is, however, open to question when applied to dynamic processes in ferromagnetic conductors.

Conflicting experimental evidence is provided by measurements carried out at the Armour Research Foundation of Illinois Institute of Technology on the dependence of the switching time on grain size in polycrystalline, randomly oriented, iron tapes of 2-mil thickness. Two tapes were prepared: one with a grain count of 26×10^4 grains per inch^2 (i.e., an average grain surface dimension of 2 mils), and another with a grain count of 6×10^4 grains per inch^2 (i.e., an average grain surface dimension of 4 mils). The switching coefficients for the two tapes differed by less than 5 per cent. One may question whether a process which originates only at grain boundaries should be nearly independent of the grain size.

It may also be questioned whether the proposed nucleation process (i.e., the initial formation of the spikelike reversal domains) is a possible one. It is not certain that, in proceeding continuously from the initial domain structure of no reversal domains to one with cone-shaped reversal domains, the total domain energy decreases uniformly. The fact that such configurations have lower energy than the initial configuration does not eliminate this objection. The question is whether the nucleated configuration is accessible without surmounting an energy barrier.

Certain points regarding the derivation of equation 5 which is intended to give the eddy-current contribution to the switching coefficient, need further clarification. The physical picture is as follows: The motion of the domain walls generates eddy currents; these eddy currents, in turn, generate a magnetic field which is counter to the external field at the domain walls, thereby decreasing the effective driving field, and hence the wall velocity.

In deriving equation 5 from this physical picture, each reversal domain is treated independently as though it occurred at the center of an alloy cylinder of radius r_m. The decrease in the eddy-current-generated counter field near the surface is neglected, even though this decrease will (at least for 1-mil tapes) have such nonnegligible effect as: 1. allowing the reversal domains formed at the surface (e.g., where grain boundaries intersect the surface) to grow much more rapidly than reversal domains formed in the center of the tape; and 2. forcing the surface reversal domains to grow much faster laterally than toward the interior. In addition, the eddy currents in that (preponderant) volume of tape which is at a distance greater than r_m from the center of the reversal domain are neglected, and the eddy-current damping of the domain wall motion is completely neglected when the radius of the reversal domain exceeds r_m.

The difficulty of treating eddy-current damping for the domain geometry chosen by the authors is such that rough approximations are unavoidable; however, the approximations introduced are so drastic that further justification is needed before equation 5 can be accepted as a valid consequence of the model on which it is based.

There is as yet no satisfactory theory of the remagnetization process in thin metallic tapes.

D. R. Brown, D. A. Buck, and **N. Menyuk:** The magnetic field required for nucleation of domains of reverse magnetization, as given by Goodenough,[1] is dependent upon grain size. This dependence is of the form

$$H_n \propto 60\sigma_w - L\omega^{*2}$$

where H_n is the required nucleation field, σ_w is the domain-wall energy per unit area, ω^* is the magnetic-pole density at the grain boundary, and L is the mean grain diameter.

Since the values of σ_w and ω^* are presumably unaffected by the change in grain size, the effect of decreasing grain size is to increase the field strength required for nucleation. In polycrystalline samples, this means that at a given field strength there will be a greater inhibition of reverse nucleation within the small grains. Therefore the ratio of the number of nucleated reverse domains to the number of grains in the sample will be smaller for the small-grain sample. Hence, there is no reason for an a priori assumption that halving the grain diameter will comparably decrease the distance between nucleated domains (and hence the switching coefficient). There is no doubt that there is a grain-size dependence of the switching coefficient. However, the foregoing equation does not yield sufficient information to permit a mathematical formulation of this dependence.

The experiment performed to determine grain-size dependence of the switching coefficient raises a question. Since this experiment was conducted upon randomly oriented iron samples, the material presumably does not have a square hysteresis loop. If this is so, equations 4 and 5 are not applicable, since these equations are based on the implicit assumptions that H_0 and d are independent of the applied field. In nonsquare-loop materials, these assumptions are inadmissible. Thus, while an apparent linear relationship between applied field and inverse switching time is frequently obtainable, the physical significance of the slope is considerably more complicated, and the intercept of the line is physically meaningless. This problem is further complicated by the grain-size dependence of the coercive force.

The second point concerns the nucleation process. The energy configuration of two possible phases were considered: the first phase is a grain completely magnetized in a single direction opposed to that of the applied magnetic field; the second phase is one in which ellipsoidal reverse domains exist within the grain. The external field value at which the second phase is more stable (i.e., the configuration with reverse domains has lower energy) is taken as the nucleation field value H_n. As the discussers point out, no consideration was given to the mechanism of transfer from the higher to the lower energy configuration.

However, since it is not at all certain that an energy barrier exists or that thermal energies are not sufficient to overcome this barrier if it does exist, the burden of proof must lie with experimental evidence which either affirms or refutes this model. Considerable experimental evidence confirming this model has been presented.[1]

The discussers raise cogent points concerning the calculation of the eddy-current con-

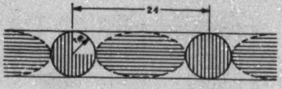


Fig. 4. Domain configuration

tribution to the switching coefficient. Let us first consider the neglect of eddy currents in that volume of the tape which is further than r_m from the center of the nucleated region of reverse magnetization. In view of this neglect, the result given in equation 5, while correct for $d = r_m$, is incomplete when $d > r_m$. To complete the calculation, it is necessary to consider the domain configuration. After the walls have moved a distance r_m, there are cylinders of completely reversed magnetization within the tape (the vertical-line area in Fig. 4). However, the magnetization reversal of the remainder of the tape must still be taken into account.

The lateral motion of the domain walls near the surface, as noted by the discussers, will be much faster than motion toward the interior. Therefore, the domain walls near the surface will rapidly spread out, and the neighboring walls will quickly meet (see dashed lines in Fig. 4). Therefore the time required must be considered for the collapse of the remaining area (the horizontal-line area in Fig. 4). This area has been assumed an ellipse of semimajor axis $d - r_m$ and semiminor axis r_m. The calculation of the time required for the collapse of this ellipse has been carried out assuming $d > 2r_m$; it was found that the result is identical with the time required for an expanding cylindrical wall to move a distance r_m. This is not surprising, since the limiting distance is r_m in both cases. Thus, if the small time required between the expanding and collapsing configurations is ignored, this model leads to a multiplying factor of 2 on the right-hand side of equation 5.

The surface effects were ignored by the authors for the sake of simplifying the model. It is probable that surface nucleation is an important factor in ultrathin tapes. It is possible to determine the effect of surface nucleation by an assumption of a domain configuration which is calculable in a standard co-ordinate system. However, any such assumption is immediately subject to the same criticism of oversimplification. In actuality there are probably both volume and surface nucleation of somewhat irregular shapes. Let us consider qualitatively what happens in this case.

The lateral motion of the Bloch walls at the surface regions of reverse magnetization will be much more rapid than their motion into the interior. The Bloch walls of the expanding regions of volume nucleation will therefore move a distance $r_1 < r_m$ and contact surface domains. This then leads to a collapsing configuration but with semiminor axis $r_2 < r_m$. From this argument it appears that the factor of 2 introduced into the right-hand side is too high. Equation 5 should read

$$S_w^e = \frac{8\alpha\pi^2 I_s r_m^2}{\rho c^2 < \cos\theta >^3} \qquad (5')$$

where $1 \leq \alpha \leq 2$.

As noted, for the sake of simplicity, these surface effects were ignored in the original calculation. In view of the limitation of the measuring technique (i.e., magnetic-field rise-time effects and switching-time definition) there was never any expectation of achieving exact numerical agreement between measured and calculated values in polycrystalline samples. However, this model does attempt to determine the parameters which affect the magnetization reversal time, and to do this quantitatively to an accuracy considerably greater than an order of magnitude.

To check the validity of this model, one of the authors has studied the switching coefficient of ultrathin 4-79 molybdenum-Permalloy tapes of four thicknesses ranging from 1/8 to 1 mil. The results verify the prediction made by this model that S_w is proportional to a constant term plus a second term which varies as the square of the thickness. The results of this experiment indicate that $\alpha / < \cos\theta >^3 = 1.3$. In addition, this experiment was carried out over a temperature range from −196 to 270 degrees centigrade. It was found that, within the limits of independent data available, the variation of the switching behavior with temperature was in accord with theoretical predictions.

REFERENCE

1. A THEORY OF DOMAIN CREATION AND COERCIVE FORCE IN POLYCRYSTALLINE FERRITES, J. B. **Goodenough.** *Physical Review*, New York, N. Y., vol. 95, 1954, pp. 917–32.

An Analogue Computer Technique Using Magnetic Amplifiers

B. E. DAVIS
NONMEMBER AIEE

I. H. SWIFT
ASSOCIATE MEMBER AIEE

THE foremost requirement to be met in the design of computers for operational use by the military services is that of providing adequate reliability of operation. Such equipment should have a mean life between failures of more than 6 months when operated under unfavorable environmental conditions. This degree of reliability is needed to insure that the equipment will inspire the user's confidence in its functional availability, and that the maintenance load under front-line conditions is kept within the capability of the services. If the designer adopts this precept, then he must limit the complexity of the equations solved in accordance with the inherent reliability of the computer components. Conventional computer elements such as servos, with their vacuum-tube amplifiers, moving parts, and potentiometers, have sufficiently low component reliability that only simple computers will meet the over-all reliability requirements. On the other hand, computers with this degree of simplicity will inadequately meet many job requirements. The best way out of this dilemma is to find new computing techniques which have a higher inherent reliability factor per unit computing operation performed.

The computing techniques described in the following sections provide a much higher reliability factor than conventional techniques, through the use of only magnetic amplifiers, metallic rectifiers, and precision resistors. These techniques have the following characteristics. 1. A computational accuracy of about 1 per cent (%) per operation, with a reasonable manufacturing cost. 2. A high ratio of computing operations performed per

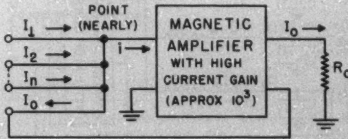


Fig. 1. Negative feedback circuit used for summing

Paper 54-389, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted March 29, 1954; made available for printing August 23, 1954.

B. E. DAVIS and I. H. SWIFT are with the Naval Ordnance Test Station, China Lake, Calif.

Jack A. Crawford, among others at the Naval Ordnance Test Station, made significant contributions to this paper.

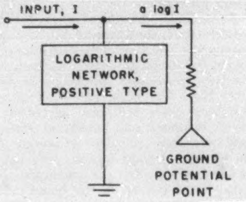
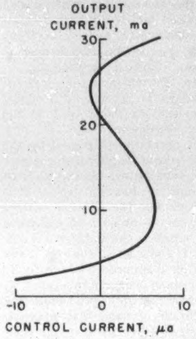

INPUT, I
a log I
LOGARITHMIC NETWORK, POSITIVE TYPE
GROUND POTENTIAL POINT

Fig. 2. Functional diagram of a logarithmic network



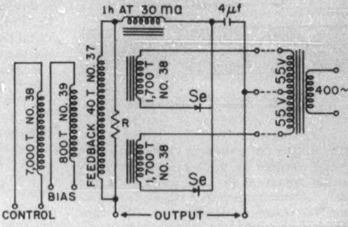

1h AT 30 ma
4μf
7,000 T NO. 38
800 T NO. 39
FEEDBACK 40 T NO. 37
R
1,700 T NO. 38
Se
1,700 T NO. 38
Se
55 v
55 v
400~
CONTROL
BIAS
OUTPUT

Fig. 4 (left). Typical transfer curve forthe magnetic amplifiers. Note that, at three points the current gain is infinite

Fig. 5 (above). Magnetic amplifier circuit. The cores are 1.125 outside diameter by 0.75 inside diameter by 0.188-inch 2-mil toroidal Deltamax. Resistor R is set for maximum gain

unit volume or weight. 3. A sufficiently short response time for a large class of applications.

Experience to date in applying these techniques to specific computer design problems indicates that they should be readily applicable to many other military and industrial problems requiring the solution of systems of nonlinear algebraic equations.

Basic Elements of Computing Method

In this scheme physical variables are represented by direct currents which flow either into or out of a ground potential point. A current flowing into ground represents a positive quantity, while a current flowing out of ground represents a negative quantity. These currents range up to 30 milliamperes (ma) in value.

Algebraic sums are obtained by a negative current feedback circuit as shown in Fig. 1. A summation of currents flowing into the input terminal of the magnetic amplifier will give

$$i = I_1 + I_2 + \ldots + I_n - I_0 \tag{1}$$

or

$$I_0 = I_1 + I_2 + \ldots + I_n - i \tag{2}$$

If the steady-state current gain of the magnetic amplifier is 10^3, then

$$i = I_0 \times 10^{-3} \tag{3}$$

and

$$I_0 = (I_1 + I_2 + \ldots + I_n)(1 + 0.001)^{-1} \tag{4}$$

The output I_0 will then be within 0.1% of its theoretical value.

Computation of products, quotients, roots, and powers is accomplished in a similar manner by using the magnetic amplifier to sum currents that are proportional to the logarithms of the input currents. These logarithmic currents are obtained by diode-resistance shaping networks, as indicated in Fig. 2. The logarithmic network indicated in Fig. 2 is a nonlinear impedance, which is so designed that the load current is proportional to the logarithm of the input current. Both positive and negative types of networks are required. Externally, the negative network looks the same as the positive network shown in Fig. 2, except that the directions of the input and output currents are reversed.

A block diagram of a logarithmic analogue computer element is shown in Fig. 3. The circuit shown here is the same as that used for summing, except that the output lead, and each input lead, has a logarithmic network applied to it. Hence, since the magnetic amplifier has a high current gain, its input current i will be small and the negative feedback action of the circuit will cause the logarithm of the output to equal the sum of the logarithms of the inputs. In Fig. 3, the output will then equal the product of the inputs, if the a's are equal. If the a's are not equal, each input will be raised to the power a_n/a_0, where a_n is the coefficient of the logarithmic current obtained from the nth input. Any of the inputs may be made to be divisors rather than multipliers by using a negative network instead of a positive one.

The Magnetic Amplifier: Design Characteristics

The principal requirement on the magnetic amplifier is that it must have a high current gain, since current drawn by its control coil will cause a computing error. The gain is therefore adjusted to be infinite at three points on its operating curve, by using sufficient positive feedback. The transfer curve of a typical magnetic amplifier is shown in Fig. 4. This curve is representative of operation in a negative feedback circuit such as that of Fig. 1 or Fig. 3. The operation of such an amplifier in either of these negative feedback circuits is best understood by first noting that the output current I_0 is determined by the nature of the feedback circuit and the values of

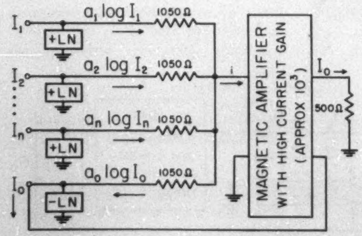

I_1
$a_1 \log I_1$
1050Ω
+LN
I_2
$a_2 \log I_2$
1050Ω
+LN
I_n
$a_n \log I_n$
1050Ω
+LN
I_0
$a_0 \log I_0$
1050Ω
-LN
i
MAGNETIC AMPLIFIER WITH HIGH CURRENT GAIN (APPROX 10^3)
I_0
500Ω

Fig. 3 (left). Block diagram of computing scheme to sum logarithmic currents

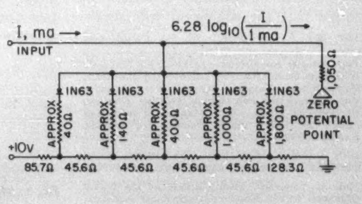

I, ma
INPUT
$6.28 \log_{10}\left(\frac{I}{1\ ma}\right)$
1N63
1N63
1N63
1N63
1N63
APPROX 40Ω
APPROX 140Ω
APPROX 400Ω
APPROX 1,000Ω
APPROX 1,800Ω
1,050Ω
ZERO POTENTIAL POINT
+10v
85.7Ω
45.6Ω
45.6Ω
45.6Ω
45.6Ω
128.3Ω

Fig. 6. Typical logarithmic network, positive type

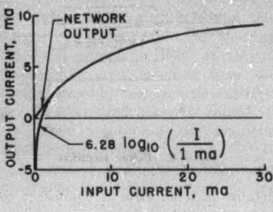


Fig. 7 (left). Illustration of the method of fitting a network output to a logarithmic curve

Fig. 8 (below). Percent error, referred to the input, for a typical network

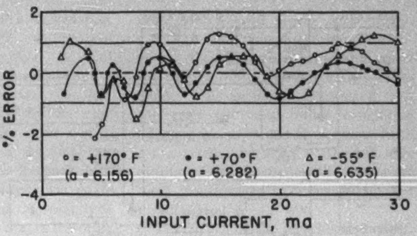
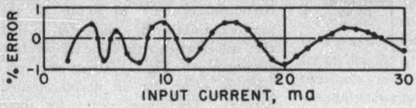


Fig. 11 (above). Network errors versus input current, for three temperatures. Each curve represents deviation from the desired equation, $I_0 = a \log (I\ ma/1\ ma)$. The values of the constant a given on the curves were chosen to minimize the errors at that temperature



the inputs. For example, in Fig. 1, if the sum of the input currents $I_1, \ldots, I_n$ is 24 ma, the output will be held to be very close to this value. Hence Fig. 4 should be viewed considering the output I_0 to be the independent variable. Then the input i is a single valued function of I_0, and its value is an indication of the deviation of the circuit operation from ideal.

The transfer characteristic of Fig. 4 is therefore meaningful only when the amplifier is used in a negative feedback circuit. The negative feedback circuit makes such an arrangement stable, and keeps the response time small. It is seen from Fig. 4 that the computing error contributed by the magnetic amplifier varies from positive to negative, depending on the output current I_0. Hence the average error will be considerably smaller than its maximum error.

Fig. 9. Package containing two summing magnetic amplifiers, with cover

Fig. 10. Logarithmic network using five germanium diodes

The error caused by the input current drawn by the magnetic amplifier may be computed for the circuit of Fig. 3. Such an analysis will readily give

$$I_0 = b^{\frac{i}{a_0}} \underset{j=1}{\overset{n}{P}} \left(I_j^{\frac{a_j}{a_0}} \right)$$

$$\cong \left[1 + \frac{i}{a_0} \log_e b \right] \underset{j=1}{\overset{n}{P}} \left(I_j^{\frac{a_j}{a_0}} \right) \qquad (1)$$

where b is the base of the logarithms, and P indicates that all quantities it operates on are to be multiplied together. The error in output current I_0 is seen to be the amount the square bracket term deviates from unity. This error is proportional to i, and independent of the values of both the output and input currents. For the design to be described in this paper, equation 1 gives an error of 1% for 27 microamperes (μa) of control current of the magnetic amplifier.

The circuit of a magnetic amplifier designed for this application is shown in Fig. 5. Fig. 9 shows a package containing two separate and identical amplifiers. The specification for this amplifier calls for an input current between plus and minus 10 μa for a constant supply voltage and under the following conditions: supply frequency, 380 to 420 cycles per second; ambient temperatures, −55 to +70 degrees centigrade; load resistance, one logarithmic network plus 500 ohms; input resistance, 800 ohms ±20%; output harmonic content less than 3%; and output currents from 2 to 30 ma. With a variation of ±10% in the supply voltage of 55 volts, the input current tolerance is doubled. This slight sensitivity to supply voltage can be compensated for by feeding the bias windings of the magnetic amplifier from the supply voltage (rectified). From the data given in the preceding paragraph, it is seen that the 10 μa maximum

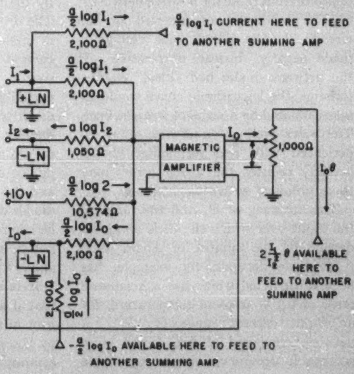

$\frac{a}{2} \log I_1 \rightarrow$
$\frac{a}{2} \log I_1$ CURRENT HERE TO FEED TO ANOTHER SUMMING AMP
2,100Ω
$I_1 \rightarrow$
$\frac{a}{2} \log I_1 \rightarrow$
+LN
2,100Ω
$I_2 \leftarrow$
$\leftarrow a \log I_2$
-LN
1,050 Ω
+10v
$\frac{a}{2} \log 2 \rightarrow$
10,574Ω
$I_0 \leftarrow$
$\frac{a}{2} \log I_0 \leftarrow$
-LN
2,100 Ω
2,100Ω
$\frac{a}{2} \log I_0$
MAGNETIC AMPLIFIER
$I_0 \rightarrow$
θ
1,000Ω
$I_0\theta$
$2\frac{I_1}{I_2^2}\theta$ AVAILABLE HERE TO FEED TO ANOTHER SUMMING AMP
$-\frac{a}{2} \log I_0$ AVAILABLE HERE TO FEED TO ANOTHER SUMMING AMP

Fig. 12 (right). Circuit illustrating some schemes used for interconnecting components to make up a computer. Also shown is the introduction of a mechanical input θ by means of a potentiometer. The symbol Δ indicates a ground potential point, the input junction of a summing amplifier

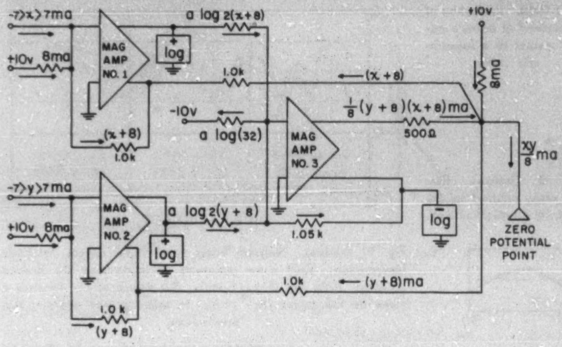


Fig. 13. Circuit for a 4-quadrant multiplier

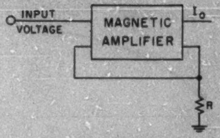


Fig. 14. Buffer amplifier

specified for the input current corresponds to a maximum computing error of 0.37%.

The Logarithmic Network: Design Characteristics

A typical circuit for a logarithmic network is shown in Fig. 6. It is made up of precision resistors and rectifiers, and its operation is as follows. When the input I is small, the rectifiers are all cut off and the output equals the input, as is shown in Fig. 7. As I is increased, the rectifiers cut in sequentially, bleeding current to ground. As shown in Fig. 7, the residual, or output current is thus made to approximate a logarithmic curve for a 15:1 range of input currents (2 to 30 ma). The network shown in Fig. 6 is termed a positive type. Negative networks are also required, as was pointed out earlier. The negative networks differ from positive ones, Fig. 6, in the following respects: 1. The directions of the input and output currents are opposite from those for positive networks; 2. The rectifiers are inverted in the circuit; 3. The 10-volt bias voltage is made negative instead of positive. If the network diodes had "ideal" characteristics, the logarithmic curve would be approximated by a series of straight lines. The more gradual on-off transition experienced with real germanium diodes actually reduces the fitting errors over those obtained with straight lines.

The accuracy of fit, and the range of the input over which the fit is obtained, depend on the number of diodes used. A 12-diode network, for example, has been constructed which has a maximum error of 0.1% at room temperature, for an input current range of 15:1. A network composed of five diodes gives acceptable accuracy, with a reasonable number of components. A typical error curve for such a 5-diode network is shown in Fig. 8. This curve shows the error, expressed as a percentage of the input, and plotted against the input current. The error is seen to be cyclic, with a maximum value of 0.8%.

The data given in Fig. 8 are typical of networks constructed according to the circuit in Fig. 6. The diodes used were *1N63* germanium, selected for having a back resistance at 75 degrees centigrade greater than 200,000 ohms at 20 volts. The values given for the resistors in series with the diodes are approximate, as they must be individually matched to the diodes. Such a network is shown in Fig. 10.

The temperature errors of a network are caused by variation of diode characteristics. A change in forward resistance of the diodes causes a change in slope of the line segments used in fitting the logarithmic curve. This factor is not very large, however, as its effect is diluted by the resistors in series with the diodes. The drop in back resistance of the diodes at high temperatures permits an error current to flow from the 10-volt reference source into the load. This error is negligible at the highest output current, as all of the diodes are biased to be conducting. At low output currents, however, this error is appreciable, as all the diodes are biased to be nonconducting, and the excess back current from all the diodes causes an error. This effect must be compensated for, if the performance is not to be degraded at high temperatures and low output currents.

It is important, however, to recognize that if a network changes with temperature in such a way that its output is off by the same percentage at all inputs, no computing error will result. This is because such an error is equivalent to computing with a different base for the logarithms. It is necessary under such circumstances only to locate all the networks together in a computer, so that they will experience the same temperature environment. Fig. 11 shows the temperature errors for an uncompensated network. The three curves, for three temperatures, each represent the error in fitting the most appropriate logarithmic curve for that temperature. It is seen that, except for low outputs at high temperatures, the errors are less than 1.5%. Such errors as these are tolerable for some computers, particularly those where input currents less than 6 ma are seldom used.

Miscellaneous Supplementary Circuits

The two basic components previously described, the logarithmic network, and the summing amplifier, have sufficient versatility that they can be interconnected to solve rather complicated systems of equations. For this, however, the designer must utilize a number of special circuits. The more important of these are described in the following.

Fig. 12 illustrates some of the circuits used in interconnecting components to form a computer. In this example, half of the output of the logarithmic network, operating on the input I_1, is fed to the summing amplifier, as shown. The other half of this current is then available to feed to another summing circuit. The same arrangement is used to make available the logarithm of the output current, I_0. By using this scheme, the logarithm of a quantity need be taken only once, and this current-splitting circuit is used to supply the logarithm to other summing amplifiers. Two points that need to be kept in mind in applying this method are: 1. The parallel combination of the current splitting resistors must equal the design load value for the networks (1,050 ohms for the networks described); and 2. The input terminal of a summing amplifier is within a few millivolts of ground potential, and for practical purposes can be considered as ground. It should be noted that a price in accuracy is paid for employment of this scheme. Equation 1 shows that splitting the load current in half, as shown, doubles the error resulting from the current drawn by the magnetic amplifier. On the other hand, the error caused by the input voltage of the magnetic amplifier is decreased.

A constant multiplier is introduced by feeding a constant current into the summing junction of the amplifier. For example, the output is multiplied by a factor of two by a resistor from a 10-volt reference source (see Fig. 12).

A mechanical input θ is introduced by the simple means shown in the output circuit of Fig. 12. The potentiometer is varied by the displacement θ, and the output is clearly zero when θ is zero. The output when the potentiometer is at the top is equal to I_0 and, for intermediate positions, is exactly proportional to the resistance from the arm of the potentiometer to ground.

The logarithmic method as described requires that the inputs always be of one sign. This restriction can usually be overcome by some means. A 4-quadrant multiplier, for example, can be constructed as shown in Fig. 13. In this circuit, a constant is added to each of the two inputs, so that the sum will always be positive. After multiplying, the unwanted product terms are then subtracted out. Note that the cost of no. 1 and no. 2 magnetic amplifiers might not be chargeable to this circuit as, in a larger computer, they will normally perform other functions too.

Often an electrical input to a computer is from a relatively high impedance source. In these cases the magnetic amplifier previously described can be used as a buffer amplifier, as shown in Fig. 14. This circuit supplies a linear output up to 30 ma, and the current drawn from the source will be between plus and minus 10 μa.

In some applications, it is desired that a relay be closed when one variable becomes larger than another. A summing amplifier can be used as an accurate comparator for this purpose. The two variables are represented by currents, one flowing into and the other out of the summing junction of the amplifier. Then, with no negative feedback connection, the amplifier output will go to its maximum value when the magnitude of the one input exceeds that of the other by more than 10 μa.

Constant, low impedance sources of both plus and minus 10 volts are required for biasing the networks, and introducing constants into the computer. These are satisfactorily provided by magnetic amplifiers similar to those described, but larger. Two of these are used in negative feedback circuits to compare their own load voltage with that of a low-power constant-voltage reference source. Since the magnetic amplifier is fully actuated by 20 microvolts on its control coil, the output is held equal to the reference source, to that tolerance.

Discussion

The preceding sections have described the two basic computing components shown in Figs. 9 and 10. These are the summing amplifier, and the logarithmic network (positive and negative types). These units, together with a plus and minus 10-volt reference source, can be interconnected to solve systems of nonlinear algebraic equations. Trigonometric and other reasonable functions can also be introduced by shaping networks similar to the logarithmic ones described. The end result is a computer without vacuum tubes which has a high density of equations solved per unit volume; a high degree of reliability if attention is paid to manufacturing detail; low power consumption and heat dissipation; and a good degree of accuracy.

As a result of using logarithms for computing, the sources of error tend to affect the output by a fixed percentage of its value. This should be contrasted with the errors of most other computing schemes, which tend to have errors which are a given percentage of the full-scale output value. Consequently, errors are most often quoted in terms of per cent of full scale.

The errors of the logarithmic scheme described depend upon the particular circuit under consideration. They usually, however, are about 1% of the output value, for each operation performed by the components described.

Experience has shown that military temperature specifications can be met using germanium diodes in the shaping networks. A factor which helps considerably in this respect is the low heat dissipation properties of the magnetic amplifiers. However, the forthcoming availability of silicon junction diodes enlarges the horizons appreciably for the computing technique described. It is stimulating to consider the possibility of replacing the relatively bulky magnetic amplifiers with power junction transistors of silicon. A computer would then consist of only silicon, precision resistors, and copper wire.

In some applications, it may be desirable to use vacuum-tube amplifiers in place of the magnetic amplifiers. Such an application may be required where very high bandwidths are required (the rise time to $1/e$ for the magnetic amplifier circuit of Fig. 1 is about 0.01 second for a 400-cycle supply). The techniques described are almost directly applicable, except that the logarithmic networks must be designed to work from a voltage source, instead of a current source. Logarithms of variables are then represented by currents, and the quantities themselves by voltages.

◆

Discussion

L. A. Finzi (Carnegie Institute of Technology, Pittsburgh, Pa.): Fig. 4 indicates that the magnetic amplifier used has a transfer characteristic with multivalued output. Characteristics of this kind are obtained commonly by introducing overregenerative ampere turns on the cores, the current of these turns being, for example, proportional to the amplifier output current or voltage.

In the device described in the paper, stable operation of the circuit is obtained by the further introduction of a degenerative feedback. If both feedbacks, overregenerative and degenerative, are provided essentially by the same variable, e.g., output current, it seems that the initial overregeneration could be dispensed with and a stable amplifier used, with suitable gain and bias to start with.

This conclusion, however, may not be justified if the variables providing the two opposing kinds of feedback are basically different. It would be desirable if the authors would elaborate on this point and on the techniques and feedback circuitry they used, for this is a matter of general interest in applications of magnetic amplifiers when very high incremental gains and stable operation are desired.

B. E. Davis and I. H. Swift: It is convenient for the purpose of evaluating the error in this computing technique, to define the current gain as the ratio of output current to control current. The amplifier described is considered as only that portion of the circuitry shown in Fig. 5. It is imperative, as has been pointed out previously, that the control current remain as small as practicable. Therefore, for use in this computing application, it is desirable to have an amplifier with near-infinite current gain.

Two components of this magnetic amplifier produce positive feedback and thus increase the current gain. The way in which rectifiers are used in this circuit increases this gain from about 8 to almost 100 and the amplifier would require 300 μa to control 30 ma of output current. An additional winding on the cores, through which the output current flows, further increases this gain by adding regenerative ampere turns. Proper selection of the number of turns of this winding produces the maximum average current gain over the entire region of operation. An average current gain of about 3,000 is obtained with the components shown in Fig. 5, resulting in the control characteristic of Fig. 4. Only 10 μa of control current are required to control 30 ma of output current after both types of feedback are inserted. The nonlinearity, which is so obvious in Fig. 4, suggests that overregenerative ampere turns have been added, but this nonlinearity arises from the components which are used and not from the feedback. This nonlinearity generally is not seen in lower gain amplifiers, because it is small compared to the larger control currents. In practice, it is frequently desirable to add more turns to the positive feedback winding and to shunt the winding with a resistor. This resistor is selected for optimum feedback in each amplifier. The use of this amplifier in a summing circuit requires that negative feedback be used. This feedback ensures that the amplifier will be stable, and it improves its response characteristics.

The positive feedback winding can be eliminated from the amplifier and the negative feedback around it reduced, so that the current gain of the whole circuit remains the same, an approach which is chiefly of academic interest, since the lower gain of the amplifier offers many detrimental characteristics to this computing technique. Most important of these are: 1. the increased interaction between input signals; 2. the error produced in the output of the logarithmic network by not connecting its load resistor to a near-ground potential point; and 3. the dependence of the external negative feedback circuitry upon the characteristics of the components within the amplifier, such as core materials and rectifiers. The positive feedback winding offers a simple and inexpensive method by which high-gain operational magnetic amplifiers can be built.

Some Applications of Semiconductor Devices in the Feedback Loop of Regulated Metallic Rectifiers

B. H. HAMILTON
ASSOCIATE MEMBER AIEE

TREMENDOUS advances have been made in semiconductor technology during the past 5 years. One of the many fields to benefit from this progress has been power rectification. The extended-area, germanium-junction diode is now a reality. The possibility of using silicon instead of germanium to provide greater freedom from ambient temperature limitations is being investigated. These new rectifying elements have improved the size, weight, efficiency, and reliability of rectifier equipments. To keep pace with the rectifying elements, there is need for corresponding improvements in regulating circuitry. It is the author's belief that this challenge will be met by the wider application of semiconductor devices in regulating circuits. Low-level junction transistors and low-leakage junction diodes are now available. Junction transistors with higher power ratings and junction diodes suitable for use as voltage standards are being developed. It is the purpose of this paper to present a few basic circuits illustrating the use of these devices to perform various functions in the feedback loop of a regulated rectifier. The type of regulated rectifier chosen for example is one suitable for charging and floating telephone central office storage batteries. As a starting point, the essential functions in the feedback loop of such a rectifier will be considered.

Block Diagram

A battery charging rectifier must be both voltage- and current-regulated. To obtain maximum life from the batteries, the output voltage must be held constant within ±1/2 per cent for any load current within the rating of the rectifier. To protect the rectifier when connected to discharged batteries, the load current must not be permitted to exceed the rating. Therefore, the rectifier must have an output voltage-current characteristic of the type shown in Fig. 1. The block diagram of a regulating system which accomplishes this is given in Fig. 2. In Fig. 2, the a-c input passes through a transformer to obtain the desired input voltage to the magnetic amplifier and rectifier. The magnetic amplifier is the power stage of the regulating circuit and controls the rectified voltage. The rectified direct current is filtered, passed through a current-sensing element, and is finally connected to the battery and associated load. Two voltage signals are fed into the gate circuit, V_V and V_I.

The voltage V_V is a portion of the battery voltage V_0. For output currents less than full load, the gate circuit admits this voltage to the error detector making V_R equal to V_V. The error detector compares V_R to the voltage standard V_S and amplifies the difference or error signal. The output of the error detector, in the circuit to be presented, is a current signal. This current is amplified by the current amplifier to provide driving current, for the magnetomotive-force-controlled magnetic amplifier. The magnetic amplifier in turn controls the battery voltage V_0 and completes the negative feedback loop. Under this condition, the changes in battery voltage caused by line voltage, line frequency, and load variations are re-

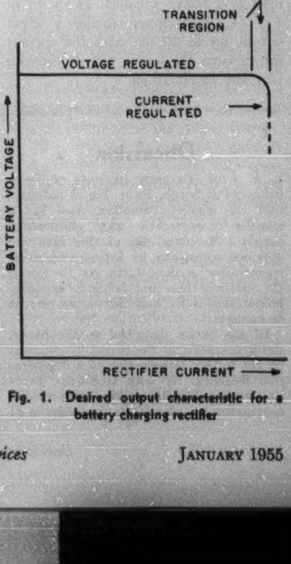


Fig. 1. Desired output characteristic for a battery charging rectifier

---

Paper 54-418, recommended by the AIEE Metallic Rectifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted July 20, 1954; made available for printing August 20, 1954.

B. H. Hamilton is with the Bell Telephone Laboratories, Inc., New York, N. Y.

The author would like to acknowledge the valuable assistance of his colleagues at Bell Telephone Laboratories.

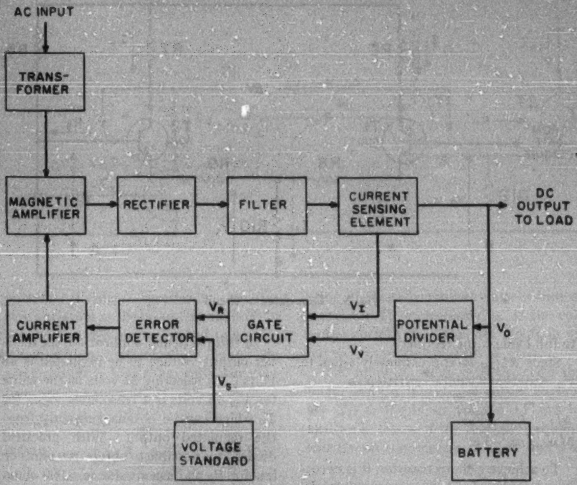


Fig. 2. Block diagram

duced by the gain of the feedback loop. The other input to the gate circuit is the voltage V_I which is proportional to the load current on the rectifier. This voltage is adjusted to be equal to V_V when the load current reaches full load. When V_I is less than V_V, corresponding to any load less than full load, the gate circuit blocks V_I from the error detector. If an attempt is made to overload the rectifier, V_I is admitted to the error detector making V_R equal to V_I. With V_R now representative of output current and tending to increase, the error detector acts to decrease the battery voltage. With V_V less than V_R, V_V is blocked by the gate circuit and the rectifier is entirely current regulated.

Semiconductor devices can be profitably applied in six of the boxes in the block diagram: the rectifier, the current-sensing element, the voltage standard, the gate circuit, the error detector, and the current amplifier. In the case of the rectifier, the major function is to convert the a-c input to a d-c output. Another requirement for the rectifier is low reverse leakage current to facilitate the use of a high gain magnetic amplifier. Large-area germanium diodes have been found well adapted to this application. The current-sensing element can take many forms. A common type of current-sensing element is an ammeter shunt. When an ammeter shunt is used, the voltage must be amplified to obtain the required voltage input to the gate circuit. A more attractive type of current-sensing element is a simple saturable reactor using high-quality core materials. The d-c input to the saturable reactor can be a single turn of the rectifier output lead. The a-c output current of the saturable reactor is approximately a square wave with amplitude equal to $2(N_c/N_g)I_L$, where

N_c = turns of the rectifier output lead
N_g = turns in the a-c winding of the reactor
I_L = d-c output current of the rectifier

To obtain a direct voltage proportional to the rectifier output current, it is necessary only to rectify the a-c output current of the saturable reactor and pass the rectified current through a resistor. The reverse leakage currents of the rectifying elements critically affect the stability of the point of transfer to current regulation. If the leakage current is not negligibly small or extremely stable, the ratio of V_I to the a-c output current of the saturable reactor will be subject to drift. Here, again, germanium- or silicon-junction diodes solve the problem by virtue of their low reverse leakage current.

With the foregoing brief treatment of the operation of the magnetic amplifier and current-sensing element as a background, the gate circuit, voltage standard, error detector, and current amplifier will be considered in detail in the succeeding sections.

Gate Circuit

As discussed in the foregoing, the function of the gate circuit is to transmit the larger of the two input voltages V_V and V_I to the output V_R. It has, therefore, a simple logic function which can be obtained with two diodes and a biasing resistor, connected as shown in Fig. 3. The higher of the two voltages V_V and V_I supplies the bias current I_b. If V_V is highest, e.g., $D1$ is biased in the forward direction and $D2$ is biased in the reverse direction. If the diodes are ideal (zero forward voltage drop and zero reverse current), then V_R is exactly equal to V_V and completely independent of V_I. Under this condition, the rectifier is voltage-regulated. As the rectifier load current increases, V_I increases, becoming equal to V_V at full load. If an attempt is made to further load the rectifier, $D2$ conducts and raises the voltage V_R. The feedback loop responds by lowering the battery voltage V_0 and, consequently, the voltage V_V. Diode $D1$ then becomes biased in

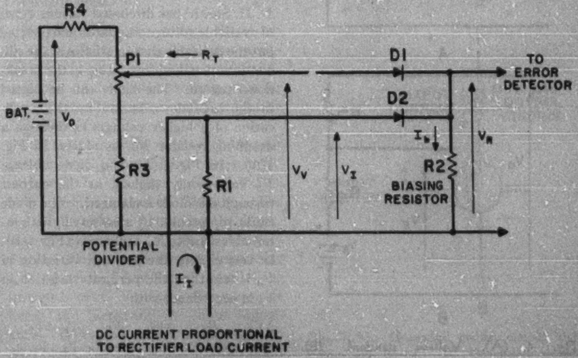


Fig. 3. Gate circuit

the reverse direction. Under this condition, V_R is equal to V_I and independent of V_V, and the rectifier is current-regulated. An important consideration is the sharpness of the transfer from voltage to current regulation using practical diodes.

In transferring from voltage to current regulation there is a transition region, indicated in Fig. 1, where the rectifier is partially voltage-regulated and partially current-regulated at the same time. In some applications, a narrow transition region is desirable. The transition to current regulation begins when the current I_I raises the voltage V_I to the point where the diode $D2$ begins to conduct appreciable current. The transition is completed when the diode $D2$ is conducting essentially all of the bias current I_b. In completing the transfer, I_I has had to increase an additional amount ΔI_I to supply the bias current and additional voltage drop ΔV_D across $D2$. These quantities are related by

$$\Delta I_I = I_b + \frac{\Delta V_D}{R1} \qquad (1)$$

The per-unit increase in rectifier load current during the transfer is

$$\frac{\Delta I_I}{I_{IM}} = \frac{I_b}{I_{IM}} + \frac{\Delta V_D}{R1 I_{IM}} \qquad (2)$$

where I_{IM} is the value of I_I corresponding

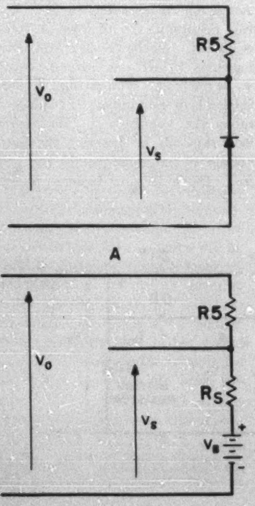


Fig. 4 (A). Voltage standard. (B). Equivalent circuit

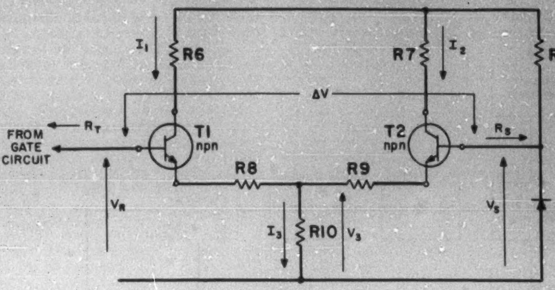


Fig. 5. Error detector circuit

to full load on the rectifier.

Since $R1 I_{IM}$ is approximately equal to V_R, equation 2 can be rewritten as

$$\frac{\Delta I_I}{\Delta I_M} = \frac{I_b}{I_{IM}} + \frac{\Delta V_D}{V_R} \qquad (3)$$

To achieve a sharp transfer, it is necessary to make I_{IM} large compared to I_b, and V_R large compared to ΔV_D. The quantities I_b and ΔV_D are governed by the characteristics of the diode. For some of the newly developed silicon diodes, practical values of I_b and ΔV_D are

$I_b = 1$ milliampere

$\Delta V_D = 0.3$ volt

Using diodes of this type, and making

$I_M = 100$ milliamperes

$V_R = 30$ volts

the transition region becomes 2 per cent.

The Voltage Standard

The precision obtained in regulating the rectifier voltage or current is limited by the stability of the voltage standard. D. H. Smith[1] has discussed various types of available voltage standards and pointed out the desirable characteristics of the silicon-junction diode operating in the breakdown region. The diode can be biased in the breakdown region by the application of a higher voltage V_0 through a dropping resistor $R5$, as shown in Fig. 4(A). In Fig. 4(A), the diode voltage V_S varies only slightly as the current through the diode is changed. The diode can be represented by a battery V_B with internal resistance R_S, as shown in Fig. 4(B). In this circuit, the per-unit variation in V_S is less than the per-unit variation in V_0 in accordance with

$$\frac{\Delta V_S}{V_S} = \frac{R_S}{R5+R_S}\,\frac{\Delta V_0}{V_S} = \frac{R_S}{R5+R_S}\,\frac{V_0}{V_S}\,\frac{\Delta V_0}{V_0} \qquad (4)$$

In this particular application, the rectifier output voltage to be regulated is 48 volts. By selecting 24 volts as the value for V_S, 48 volts is a reasonable value for V_0 which can be obtained directly from the regulated output. With practical diodes, it is possible to obtain 100 ohms or less for R_S and use a value of 5,000 ohms for $R5$. With these constants, the percent variation of V_S can be evaluated when the rectifier is voltage-regulated within ±1/2 per cent, as mentioned earlier. The calculated per-cent variation in V_S is ±0.02 per cent, which degrades the precision of regulation by a negligible amount. However, when the rectifier is current-regulated, the output voltage V_0 is subject to wider variation, and the precision of current regulation is thereby lessened.

Another possible source of variation in the voltage standard is temperature. The voltage V_B in the equivalent circuit of Fig. 4(B) has, in general, a temperature coefficient. At the present early stage of development of the silicon-junction diode, the ultimate control to be realized over the magnitude and uniformity of the temperature coefficient is not yet fully known. Temperature coefficients of 0.03 per cent ± 0.01 per cent per degree centigrade have been observed for diodes with breakdown voltages in the range of 5 to 6 volts. It is further feasible to construct a voltage standard using more than one diode to obtain a zero nominal temperature coefficient; see reference 1.

The Error Detector

The error detector must compare the output voltage of the gate circuit V_R to the reference voltage V_S and provide a signal proportional to the difference. The power available to the input of the error detector is limited by the internal resistance of the voltage standard and the Thévenin equivalent resistance of the

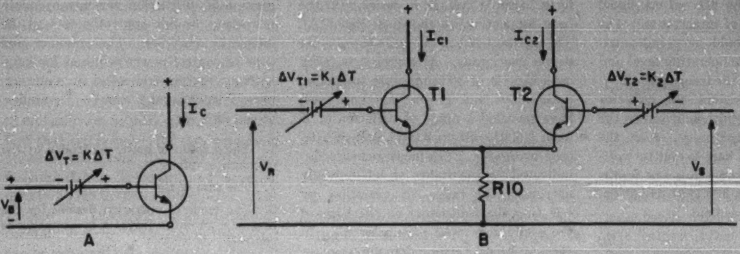


Fig. 6 (A) Grounded emitter circuit. (B). Push-pull circuit

potential divider. Specifically, the current variation at the input terminals of the error detector must be limited to a value which will not appreciably "load" the error voltage. Mathematically, this relation can be expressed as

$$\Delta I_0(R_T+R_S) \ll \Delta V \tag{5}$$

where

ΔI_0 = variation in current at the input terminals of the error detector
ΔV = permissible error voltage
R_T = Thévenin equivalent resistance of the potential divider
R_S = internal resistance of the voltage standard

Equation 5 can be written to give the required input resistance of the error detector

$$\text{Input Resistance} = \frac{\Delta V}{\Delta I_0} \gg (R_T+R_S) \tag{6}$$

In the example of a 48-volt rectifier regulated within ±1/2 per cent and using the diode voltage standard described earlier, ΔV is 0.24 volt and R_S is 100 ohms.

Using an assumed value of 100 ohms for R_T, equation 6 shows that the input resistance of the error detector must be large compared to 200 ohms.

Two 50-milliwatt junction transistors can be used in the circuit shown in Fig. 5 to provide error detection and amplification, meeting the requirements outlined in the foregoing. In Fig. 5, the transistors are emitter-coupled and the error voltage $\Delta V = V_R - V_S$ is applied between the base terminals. The output signal is the difference between the collector currents I_1 and I_2. The input resistance of the push-pull current amplifier, to be described later, is represented by the resistors *R*6 and *R*7.

To simplify the explanation of the operation of this circuit, assume *R*8 and *R*9 are equal to zero. Then, in response to an increase in the voltage V_R, the base to the emitter voltage bias of transistor *T*1 is increased. This causes the collector current I_1 and the emitter current of *T*1 to increase. Through the emitter coupling, the base to emitter voltage bias of transistor *T*2 is decreased as is the collector current I_2. The net result is an increase in output current (I_1-I_2).

The input impedance to the error detector and the ratio G_M of output current (I_1-I_2) to the input voltage ΔV can be written in terms of the transistor equivalent circuit parameters r_e, r_b, r_c, and α as

$$\text{input impedance} \approx 2\left(r_b+\frac{r_e}{1-\alpha}\right) \tag{7}$$

$$G_M = \frac{I_1-I_2}{\Delta V} \approx \frac{\alpha}{r_e+(1-\alpha)r_b} \tag{8}$$

assuming

$$\frac{r_e}{1-\alpha}+r_b \ll r_c \tag{9}$$

Typical values for the transistor parameters are

$r_e = 8$ ohms (for emitter current = 3 mils)
$r_b = 250$ ohms
$\alpha = 0.95$
$r_c = 5\times10^6$ ohms

Substitution of these values in equations 7 and 8 gives the following values for input impedance and G_M

input impedance = 820 ohms
$G_M = 0.046$ mho

By multiplying the voltage signal ΔV by the G_M factor, an output current signal (I_1-I_2) is found to be 11.0 milliamperes. A desirable operating range for the collector currents I_1 and I_2 might be 2 to 4 milliamperes.

The total operating range of the output current signal would then be 4 milliamperes. In this case a G_M of 0.0166 mho is desired instead of 0.046 mho. The G_M can be reduced by adding the degenerative resistors *R*8 and *R*9, as shown in Fig. 5. This is equivalent to increasing r_e in equations 7 and 8, and provides the further benefits of making the gain more linear and increasing the input impedance. If *R*8 and *R*9 are made equal to 36.5 ohms, the desired G_M (0.0166 mho) is obtained and the input resistance is increased to 2,280 ohms.

An appropriate range of collector voltage variation is 10 volts. This limits the input resistances of the current amplifier *R*6 and *R*7 to 5,000 ohms. The output power range of the error detector is, then, 4.0×10^{-2} watts. With the input resistance of 2,280 ohms, the input power range is 1.26×10^{-5} watts. The power gain of the error detector is 35 decibels.

The only remaining question as regards the suitability of the transistor error detector is: What about drift effects with ambient temperature variations? There are two types of output drift which could affect the operation of the error detector.

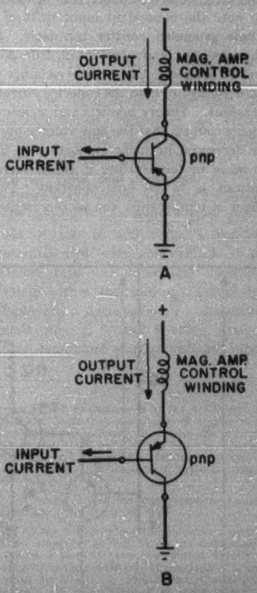


Fig. 7(A). Grounded emitter circuit, high output impedance. (B). Grounded collector circuit, low output impedance

Any variation in the net output signal (I_1-I_2) is certainly of importance. The quantity (I_1+I_2) could vary independently and change the operating level and power dissipation of the transistor.

In junction transistors, the collector current is related to emitter current by the factor α which is near unity. Also, the range of variation of base to emitter voltage in the working range is a few tenths of a volt. Thus, to a first approximation, the following relations hold

$$I_3 \approx I_1+I_2 \tag{10}$$

$$V_3 \approx V_S \tag{11}$$

Since V_3 is equal to $I_3\,R10$, equations 10 and 11 can be used to obtain the further approximation

$$I_1+I_2 \approx \frac{V_S}{R10} \tag{12}$$

Thus, the quantity (I_1+I_2) is dependent primarily on parameters external to the transistor and will be affected only to a secondary extent by changes in transistor parameters such as the temperature-sensitive saturation current I_{c0}.

The drift effect of temperature on the output signal is most easily considered in terms of the equivalent drift at the input. Before attempting to analyze the push-pull error detector circuit, it is beneficial to note the equivalent input drift of a single grounded emitter transistor. In this case, the equivalent input drift ΔV_T will be defined as the necessary change in base voltage to hold the collector current constant. Theory predicts and measurements confirm that the base voltage must be decreased in linear proportion to the increase in temperature if the collector current is to be held constant. This effect of temperature can be represented by a variable voltage in series with the base connection, as shown in Fig. 6(A). The constant K has the units of volts per degree centigrade. Measurements taken on a sample of 30 transistors (including both alloy- and grown junction-types) have revealed a range of variation of K from 0.0021 volt to 0.0032 volt per degree centigrade. The measurements also indicated the possibility of substantially narrowing the range of variation by selecting the transistors on the basis of the easily measured parameter h_{11}. In the push-pull error detector circuit, the temperature coefficients $K1$ and $K2$ are subtractive, as shown in Fig. 6(B). The per-cent drift in output voltage caused by the temperature coefficients of transistors in the error detector is given by

$$E=\frac{K_1-K_2}{V_S}\times 100 \tag{13}$$

where

E = the per-cent drift in the rectifier output voltage per degree centigrade
K_1 = temperature coefficient of transistor $T1$
K_2 = temperature coefficient of transistor $T2$
V_S = voltage of the standard

Even with badly matched transistors ($K_1=0.0021$, $K_2=0.0032$), the drift in the regulated voltage, from equation 13, is 0.0046 per cent per degree centigrade.

The foregoing analysis has demonstrated the suitability of the transistor error detector circuit from the standpoints of input impedance, gain, and stability with temperature.

The Current Amplifier

The final link in the feedback circuit is the current amplifier. In the block diagram under discussion, the current amplifier must provide control power to the magnetic amplifier. The amount and type of control power required for magnetomotive-force-controlled magnetic amplifiers depends on a variety of considerations such as:

1. The amount of power controlled by the magnetic amplifier.
2. The gain of the magnetic amplifier.
3. The mode of operation desired for the magnetic amplifier (high or low impedance control).
4. The voltage or current induced in the magnetic amplifier control windings.

The first two considerations determine the required output power of the current amplifier. If this power is not too great, an all-transistor current amplifier can be used. Otherwise, it may be necessary to use a magnetic amplifier in the output stage. The ultimate power rating obtainable from practical transistors is yet to be determined. Transistors using special cooling methods have been reported capable of dissipating up to 20 watts. Transistors are well adapted to provide current amplification with either low or high output impedance. To obtain high output impedance, the grounded emitter circuit of Fig. 7(A) is suggested. With this circuit output impedances up to tens of thousands of ohms can be obtained. For low output impedance, the grounded collector circuit of Fig. 7(B) is preferable. In this case, output impedance of tens of ohms can be obtained. It is not intended to cover the field of magnetic amplifier control with transistor circuits. Instead, a push-pull transistor current amplifier which the author has found particularly flexible will be presented.

In the current amplifier circuit shown in Fig. 8, four pnp junction transis-

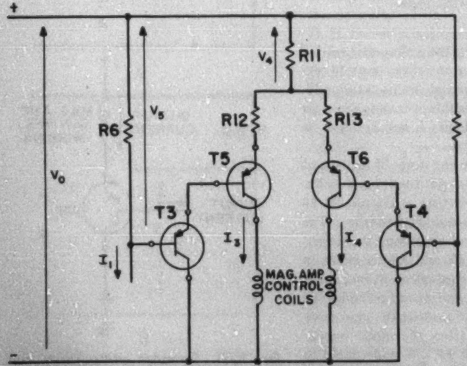


Fig. 8. Push-pull current amplifier

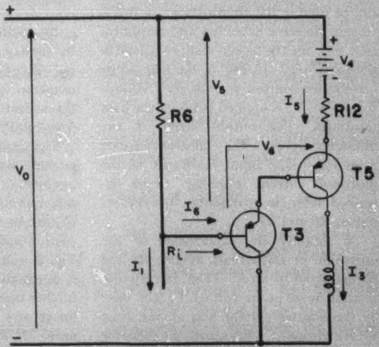


Fig. 9. Half of push-pull current amplifier

tors are used to provide a current amplification of 100. The input currents I_1 and I_2 are the output currents of the error detector circuit in Fig. 5 which vary from 2 to 4 mils in push-pull fashion. The output currents I_3 and I_4 vary from 0 to 200 in milliamperes in push-pull. To the extent that the I_3 and I_4 do vary equally in opposite directions, the current through resistor *R*11 is constant and approximately equal to 200 milliamperes. The voltage V_4, therefore, is approximately constant and the two halves of the circuit can be analyzed independently. The left half of the circuit has been redrawn in Fig. 9 with *R*11 replaced by a constant voltage V_4.

In Fig. 9, the driving current required at the base of transistor *T*3 is given by

$$\Delta I_6 = \frac{\Delta I_3}{G_1 G_2} \tag{14}$$

where

ΔI_6 = driving current
ΔI_3 = range of variation of output current, 200 milliamperes
G_1 = current gain of transistor *T*3
G_2 = current gain of transistor *T*5

A practical value for the product of G_1 and G_2 is 1,000. The driving current ΔI_6 calculated from equation 13 is then 0.2 milliampere. Since ΔI_6 is small compared to the 2-milliampere variation in I_1, the following approximation is obtained

$$\Delta V_5 \approx \Delta I_1\, R6 \tag{15}$$

A desired value for *R*6 is 5,000 ohms, as indicated in the analysis of the error detector circuit. The variation in V_5, then, becomes 10 volts. The normal variation of the voltage between the base of *T*3 and the emitter of *T*5 is negligible compared to 10 volts. This leads to the approximation that

$$\Delta V_5 \approx \Delta I_3\, R12 \tag{16}$$

By combining equations 15 and 16, equation 17 results

$$\Delta I_1\, R6 \approx \Delta I_3\, R12 \tag{17}$$

Since the emitter and collector currents of *T*5 are approximately equal, equation 17 can be rewritten as

$$\frac{\Delta I_3}{\Delta I_1} \approx \frac{R6}{R12} \tag{18}$$

Equation 18 shows that the current gain of the amplifiers is controlled, to a first approximation, only by the passive components *R*6 and *R*12. It is, therefore stabilized against variations in the transistor parameters with temperature.

Conclusion

The basic circuits presented here were not intended to be all-inclusive. Rather, they were intended to illustrate the flexibility of low-level semiconductor devices in the feedback loop of comparatively large power rectifiers. One of the attractive features is that the transistor amplifiers can use the regulated output voltage as a power supply in many cases where vacuum tubes would require a separate auxiliary power supply. It is believed that this application of the new semiconductor devices will supplement the progress being made with power diodes and lead to further improvements in the size, weight, efficiency, and reliability of regulated rectifier equipment.

Reference

1. The Suitability of the Silicon Alloy Junction Diode as a Reference Standard in Regulated Metallic Rectifier Circuits, D. H. Smith. *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1955 section), pp. 645–51.

The Suitability of the Silicon Alloy Junction Diode as a Reference Standard in Regulated Metallic Rectifier Circuits

D. H. SMITH
MEMBER AIEE

Synopsis: Regulated metallic rectifiers have been used throughout the Bell System to provide talking current, to power relays, switches, and vacuum-tube filaments; and to supply vacuum-tube plate potentials. With the advent of semiconductor devices such as transistors and junction diodes, the regulated metallic rectifier is assuming added importance. However, the lower voltages generally required by these semiconductor amplifier and switching devices have required a re-evaluation of the regulating principles used in metallic rectifiers. It has been found possible to replace previous vacuum-tube and magnetic-amplifier regulating circuits with equivalent and usually more compact and efficient semiconductor apparatus. This paper discusses the suitability of silicon-alloy junction diodes having very sharp reverse breakdown characteristics in the reference standard portion of closed-loop feedback regulating circuits of metallic rectifiers.

PREVIOUS papers by the author and others have discussed the use of metallic rectifiers in telephone power plants,[1] the automatic regulation of metallic rectifiers by vacuum-tube electronic control (unpublished), and the automatic regulation of metallic rectifiers by magnetic amplifier control.[2,3] With the advent of semiconductor devices such as transistors and junction diodes it has become necessary to re-evaluate known methods of automatic regulation of metallic rectifiers in the light of this new art.

First, the new semiconductor apparatus requires much lower electrode potentials than the vacuum-tube electronic and electomechanical analogues. A typical common emitter transistor amplifier utilizes collector potentials between 6 and 20 volts, and base bias potential between 2 and 6 volts, whereas a vacuum-tube amplifier may require plate potentials of 130 to 400 volts, a filament potential of 6 volts, and perhaps 20- to 50-volt bias. At least one computer design within the author's knowledge, employing semiconductor apparatus, needs only 2, 6, and 8 volts, whereas other computers employ vacuum tubes,

---

Paper 54-417, recommended by the AIEE Metallic Rectifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted July 20, 1954; made available for printing August 20, 1954.

D. H. Smith is with the Bell Telephone Laboratories, Inc., New York, N. Y.

The author acknowledges the valuable assistance of his colleagues at Bell Telephone Laboratories. At least a score of them contributed directly or indirectly in the preparation of this paper. Special thanks are due J. R. Flegal who furnished the diodes, C. W. Van Duyne, E. E. Helin, and R. P. Drews who organized the test program, and B. H. Hamilton who contributed many valuable suggestions and ideas, and who is already planning to use the diodes as standards in some of his circuits.[7]

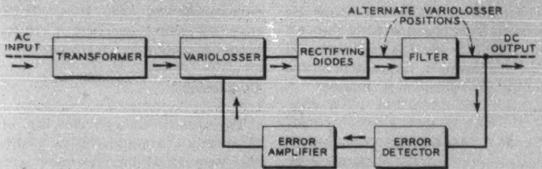


Fig. 1. Basic elements of feedback-regulated rectifier

relays, switches, and cathode-ray tubes requiring a multiplicity of high and low voltages. Thus, the first problem to be faced by the designer of regulated rectifiers is that of the generally lower load voltages and, consequently, the lower error signals with which he must deal.

Second, if the rectifier designer is to progress with the times, he must employ these new semiconductor devices in his regulating circuits where it is possible to take advantage of their compactness, efficiency, and reliability.

This paper discusses how Bell Telephone Laboratories has approached the general problem of precise power regulation at low voltages, and gives an introduction to the design of feedback regulating amplifiers employing semiconductor apparatus. Future papers will cover in more detail this latter topic.

Basic Regulating Circuit

Fig. 1 gives in block diagram form a generalized feedback-regulated rectifier circuit. The elements include a step-up or step-down transformer, a variolosser, the rectifying diodes, the filter, the error detector, and the error amplifier. Variolosser is a term borrowed from others in carrier transmission work, who use it to denote a variable gain amplifying stage which can amplify or attenuate a signal.

The Variolosser

In power regulator work, a variolosser functions usually to absorb more or less of the voltage available to the load and can function in the d-c or the a-c portion of the circuit. A typical d-c variolosser is a series vacuum tube which acts as a variable resistor in series with the output. A typical a-c variolosser is a continuously tapped autotransformer or a magnetic amplifier. Fig. 2 illustrates some forms the variolosser may take. A simple form of d-c variolosser is a variable resistor in series with the load current, which may be manually adjusted to maintain a constant voltage across the load. Series tubes and series transistors may be thought of as automatic-type variable resistors. Another form of variolosser utilizes a variable shunt resistor and a fixed series or ballast resistor. Here, too, vacuum tubes or transistors can be used as the shunt element. A grid-controlled thyratron variolosser combines a rectifying function with a gating function, as does a magnetic amplifier variolosser.

Rectifying Diodes

The past decades have seen vast improvements in rectifying diodes, from the copper oxide rectifier in the 1920's and the selenium diodes in the 1940's, to the germanium and silicon junction diodes in the 1950's. The improvements in both forward voltage drop and reverse resistance have been substantial, but detailed discussion of these is beyond the scope of this paper.

Error Detector

The error detector and error amplifier are used to complete the feedback loop as shown in Fig. 1. The rest of this paper will be devoted to a discussion of reference standards used in error detector circuits, with emphasis on a new type of reference standard which was recently developed.

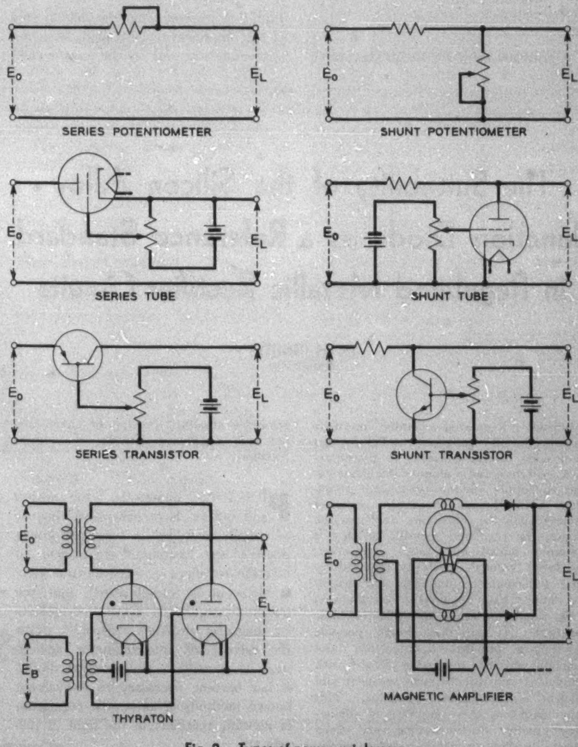


Fig. 2. Types of power variolossers

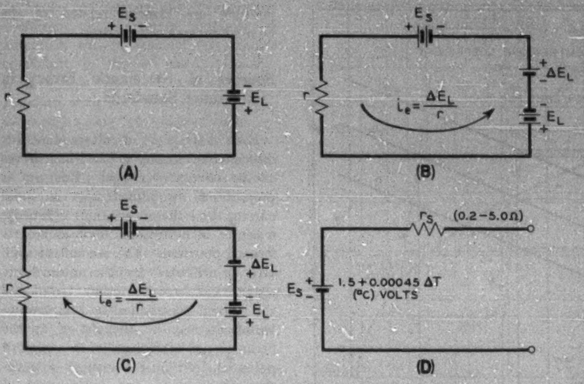


Fig. 3. Battery reference

Battery Reference

Fig. 3 illustrates the use of a battery as a reference standard in an error detector circuit. In Fig. 3(A), the desired potential of the load is E_L, the potential of the reference standard is E_s, and r is the input resistance of the error amplifier.

As long as $E_L=E_s$ no feedback current flows and no regulating action takes place. If, however, E_L increases by an amount ΔE_L, as in Fig. 3(B), the net voltage in the feedback loop is seen to be ΔE_L, and an error signal equal to $\Delta E_L/r$ flows in the input of the error amplifier. Likewise, if E_L decreases by an amount ΔE_L, as in Fig. 3(C), the net voltage in the feedback loop is again ΔE_L, but since ΔE_L now has an opposite sign, an error signal equal to $-\Delta E_L/r$ flows in the input of the error amplifier.

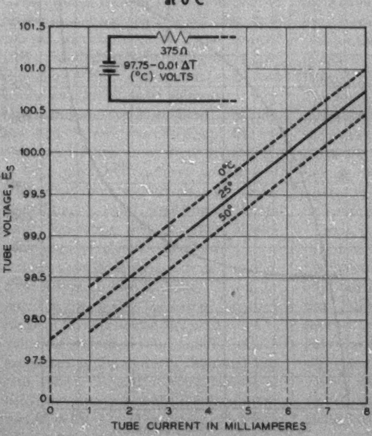
Batteries have been used extensively as reference standards in regulated rectifiers, primarily in vacuum-tube grid circuits[1] where the reference battery does not have to furnish over 100 microamperes of current. An important limitation of a battery used quite extensively in the Bell System as a standard is an internal resistance, shown in Fig. 3(D) as r_s. Although r_s may be less than 5 ohms per cell when a battery is new, the resistance increases with age and finally reaches a value which causes appreciable errors in the regulating circuit. Also to be noted is the positive temperature coefficient of the electromotive force (emf), about 0.00045 volt per degree centigrade (C). Over a 50 C range, the emf would change approximately 0.02 volt per cell. Another way of expressing the temperature coefficient is in percentage of the nominal open-circuit voltage (1.5 volts) and this works out to be +0.03 per cent per C.

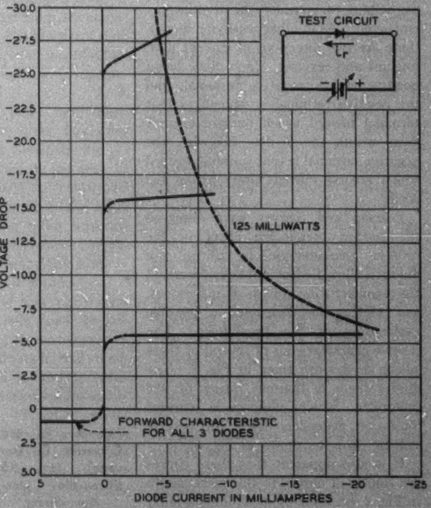
Glow-Discharge Tube Reference

Glow-discharge tubes are frequently used as reference standards.[4] The volt-ampere characteristic of a typical glow-discharge tube, designed expressly to be used as a reference is given in Fig. 4. The glow-discharge tube is seen to be equivalent to a battery having an open-circuit voltage of approximately 100 volts, an internal resistance of about 375 ohms, and a negative temperature coefficient of 0.01 per cent per C. The tube is like a battery, but it requires application of a potential higher than the reference voltage required to initiate conduction in the tube. For the tube under discussion, the maximum breakdown voltage is about 165 volts. If the regulated load voltage is less than 165 volts, a separate d-c supply must be

Fig. 4 (below). Glow-discharge tube reference

Fig. 5 (right). Silicon-alloy junction diode characteristics at 0 C





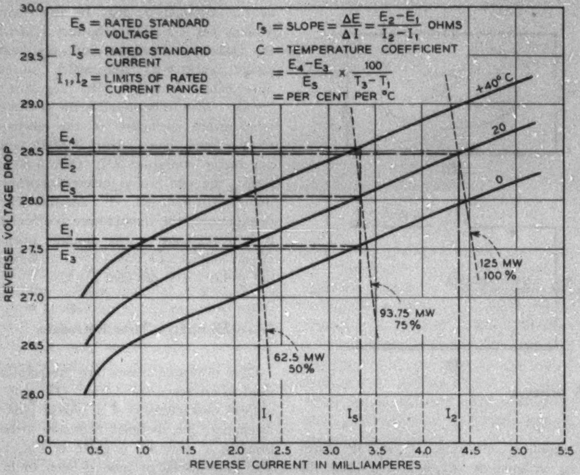


Fig. 6. Derivation of slope and temperature coefficient

provided in the regulating circuit to insure breakdown. A potentiometer circuit is usually employed to make the standard voltage nearly equal to the regulated voltage at the point of comparison.

Alloy Junction Diode Reference

Before discussing the characteristics of junction diode references, it may be helpful to consider the terminology employed. As in other rectifying diodes, there are two directions of current flow, forward and reverse. Each diode has a positive and a negative terminal, and the positive terminal is defined as that terminal toward which forward current flows within the diode. Likewise, the negative terminal is that terminal toward which reverse current flows within the diode.

Pearson and Sawyer have observed the sudden breakdown effect in the reverse characteristic in silicon *p-n* junction alloy diodes[5] and have described the manufacturing process of such diodes, and McKay has discussed a possible explanation of the breakdown.[6] The reverse characteristics of several of these silicon diodes are given in Fig. 5. The diodes have a thermal rating of 125 milliwatts, which imposes a current limitation dependent upon the reverse voltage drop, as shown in Fig. 5.

It is also apparent that these curves are similar in nature to the curve shown in Fig. 4, i.e., each diode exhibits a fairly constant voltage drop over a wide range of reverse current. Also to be noted is the almost constant forward voltage drop. The voltage drop varies with temperature and with current. The variation with current is seen to be greater for the higher voltage diodes than with the lower. It has been found that the voltage variations with temperature are likewise related to the nominal reverse voltage drop.

Sources of Systematic Errors in Reference Standards

Each of the types of voltage standards discussed has a counter emf which serves as the reference voltage. This emf is proportional to both temperature and current, and these variations constitute a source of systematic error in the reference potential. The amount of such error contributed by the internal resistance can be called "slope error" and the amount contributed by temperature variations may be thought of as the "temperature coefficient" of the standard potential. Within the author's observation, these errors are independent of each other, i.e., the slope is constant over a wide range of temperatures, and the temperature coefficient is constant at all values of current within the rating.

In general, it is desirable to minimize internal errors in a reference standard. There are circumstances where a known internal error can be used to compensate for a known error external to the standard. For instance, a given slope error can be utilized to offset a portion of the change in rectifier output voltage caused by a-c input voltage changes, and thereby lessen the amount of automatic regulation range required of the feedback circuit. Similarly, in some circuits it may be

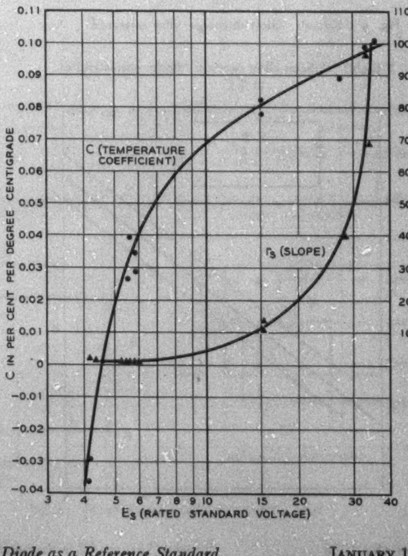


Fig. 7. R_S and C versus E_S (reverse) for 15 silicon-alloy junction diodes

desirable to balance the temperature coefficient of the error amplifier or amplifiers with a residual temperature coefficient in the standard. Usually, however, it is desired to obtain a reference standard with zero temperature coefficient and a very low value of slope.

The slope error of a standard can be determined by plotting the change of terminal voltage over a given current range in the breakdown region, and dividing the change in voltage by the change in current. This is sometimes called a-c, differential, or dynamic resistance, to differentiate it from the absolute resistance reckoned from the current and voltage axes.

The temperature error can be determined by holding the current constant at some nominal value in the breakdown region and noting the change in voltage drop at various temperatures. The temperature coefficient may then be defined as the change in voltage, divided by the nominal voltage and the change in temperature in C. This factor is then multiplied by 100 to express the coefficient in per-cent change per C.

These two parameters were evaluated for a sample of 15 experimental silicon-alloy junction diodes, having nominal reverse voltage drops between 4 and 40 volts. Curves of reverse voltage versus reverse current were plotted at 0, +20, and +40 C. The reverse current was varied in each diode from one milliampere to a value representing a power dissipation of approximately 125 milliwatts, the tentative full power rating of this type of diode.

Next, isodissipative lines were superimposed on each set of curves at power levels of 62.5, 94, and 125 milliwatts, representing 50 per cent, 75 per cent, and 100 per cent of the power rating respectively. Rated ambient temperature was chosen as +20 C, and nominal operating power was chosen as 94 milliwatts. The intersection of the 94-milliwatt line and the +20 C curve was noted, and the coordinates of this point were termed E_s and I_s; the rated reference voltage and reference current respectively. It was expected that E_s and I_s would be the design center operating point of these diodes when used as reference standards.

This process is demonstrated in Fig. 6 for a typical 28-volt diode. It will be noted in Fig. 6 that the linear operating range of this particular diode extended to values of current beyond I_1 and I_2, which are stated as the limits of the rated current range. I_1 and I_2 were chosen as being the abscissas of the points of intersection of the +20 C operating curve with the 62.5-milliwatt (mw) and 125-mw power level curves respectively. This was judged necessary because several of the test diodes departed from linear operation just below the 62.5-mw point, but it is expected that improved manufacturing techniques will permit a wider operating range in the future.

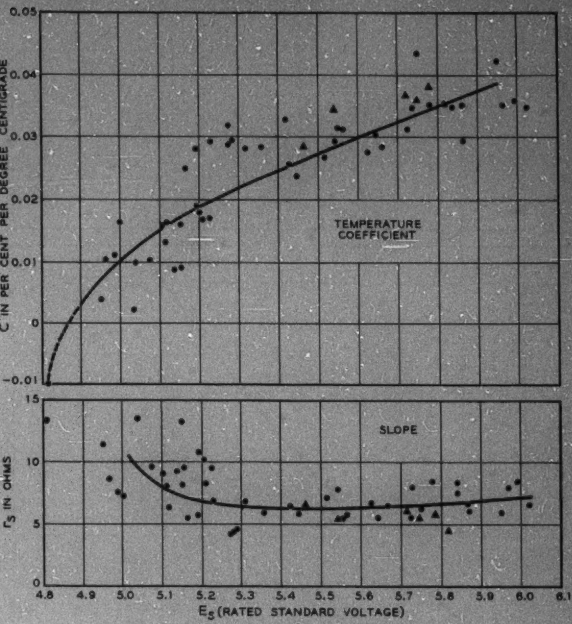


Fig. 8. R_s and C versus E_s (reverse) for 57 silicon-alloy junction diodes

After the values of r_s and C were determined for the samples, these data were plotted against E_s, as shown in Fig. 7. The temperature coefficients are seen to range from about -0.04 to $+0.10$ and the values of r_s vary between 10 and 1,000 ohms for diodes having rated standard voltages between 4 and 40 volts. Several interesting characteristics of this particular batch of diodes can be noted.

First, it would appear that a zero temperature coefficient diode could be obtained if the manufacturing processes were adjusted to produce diodes having standard voltages between 4 and 6 volts. Second, these same diodes would probably have minimum slope. Third, it would seem more feasible to make a medium voltage standard with several low-voltage diodes connected in series than with one diode. Note that a 24-volt reference standard made up of four 6-volt diodes would have a temperature coefficient of +0.035 and a slope of 40 ohms, as opposed to a single 24-volt diode which would have a temperature coefficient of +0.094 and a slope in excess of 300 ohms. A $2^1/_2$-to-1 improvement in C and a $7^1/_2$-to-1 improvement in r_s could be realized by using the 4-diode standard.

Although the exploratory work had been limited to a small sample of diodes up to this point, it appeared that there was a definite correlation between E_s, r_s and C which should be investigated further. Accordingly, a somewhat larger sample of about 60 diodes was obtained, but this time the manufacturing process was adjusted to secure breakdown voltages between the limits of 4 and 6 volts. Referring again to Fig. 7, it can be seen that the temperature coefficient curve crosses the axis at about 4.5 volts. Thus, it appeared that the range in breakdown voltage between 4 and 6 volts was of most interest. The 60 diodes were subjected to the previously described test procedure and the results of the test are plotted in Fig. 8. In addition, it was decided to perform the same test pro-

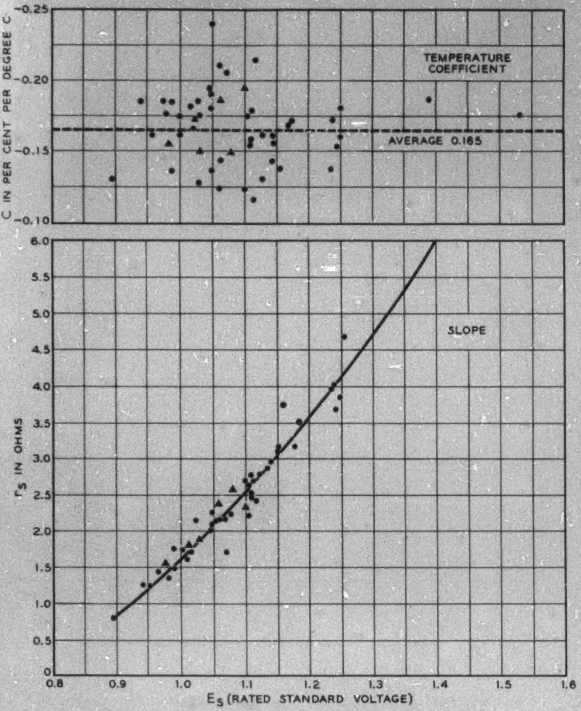


Fig. 9. R_S and C versus E_S (forward) for 57 silicon-alloy junction diodes

cedure with respect to the forward voltage drop of these particular diodes, and these data are plotted in Fig. 9.

Reverse Voltage

The data in Fig. 8 show good agreement with the behavior of the first test samples. All but two diodes had breakdown voltages between 4.9 and 6.1 volts. Most diodes in this group had slopes in the range of 4.5 to 10 ohms, and temperature coefficients between +0.0023 and +0.04. Of the two diodes remaining, one had a breakdown voltage of 4.8 volts, a slope of 13.3 ohms and a temperature coefficient of −0.01. This was the only diode having a negative value for C. It is expected that if a substantial sample of diodes had been obtained in the 4-to-4.5 volt region, most of these would have had negative values of C. The other diode had a breakdown voltage of 6.14, a slope of 17 ohms, and a temperature coefficient of +0.026.

Forward Voltage

The data in Fig. 9 indicate that the forward voltage drop may be used as a reference standard, but that C will probably always be negative, r_s will be 2 or 3 ohms, and E_s will be 1 to 1.2 volts.

Long-Time Stability and Life

It can be seen that systematic changes in the reference potential are probably closely related to the breakdown voltage and the environmental temperature. It is then valid to ask:

1. Are there any random fluctuations in the voltage?
2. Are there any time-dependent linear variations in the voltage?
3. What is the expected life?

In the case of the glow-discharge tube previously discussed, random and linear changes are about 0.1 volt and 0.2 respectively. The life is estimated at well over 2 years. The battery has no discernable random fluctuations, but local chemical action causes self-discharge and a gradual increase in the internal resistance. Eventually the open-circuit voltage falls to zero. Life expectancy as a reference is about 2 years.

The silicon junction diode is not expected to exhibit either random- or time-dependent variations in reverse or forward voltage drop when operated within its thermal rating. Several of these diodes have been tested for a continuous period in excess of 9,000 hours and neither type of fluctuation has been observed. The author believes that the physical theory explaining the properties of the silicon diode bears out this observation. Additional assurance must, of course, await more experience both in the manufacturing process and in circuit use. This is probably the place to point out that, like other semiconductor junction devices, the electrical properties of the junction are affected by moisture, but it is expected that hermetic sealing will provide adequate protection.

Conclusions

Through the judicious use of several diodes connected in series, some in the forward direction and others in the reverse direction, it should be possible to "tailor-make" a voltage standard having close to a zero temperature coefficient, a net positive temperature coefficient, or a negative temperature coefficient, as desired.

Values of slope of less than 200 ohms should be obtainable for voltage standards of about 100 volts using a series diode arrangement. For voltage standards considerably less than 100 volts, such as might be desired for a 2-, 10-, or 50-volt regulated rectifier, the diode voltage standard should compete strongly with other references for it will be physically small and have long life and a high order of stability.

The future of this type of voltage standard depends largely upon the cost, which in turn is closely related to demand. In view of the fact that the silicon junction diode has many other uses, such as in d-c "choppers," in frequency-modulation discriminators, in surge protection circuits, and as clippers, as gate valves in computer circuits, and even as low-voltage and low-current power rectifiers, it is possible that the necessary demand will be forthcoming.

In any event, it is believed that designers of regulated rectifiers should consider the use of the silicon-alloy junction diode as a reference standard. It should be most popular in circuits employing transistors, or combinations of transistors and magnetic amplifiers, where its small size, expected long life and stable characteristics will be commensurate with similar parameters of the rest of the circuit. Other papers[7,8] discuss in more detail how the diode may be employed in these circuits.

References

1. Metallic Rectifiers in Telephone Power Plants, **D. E. Trucksees.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 1464–67.

2. Automatic Regulation of Metallic Rectifiers by Magnetic Control, **D. H. Smith.** *AIEE Transactions*, vol. 71, pt. I, Jan. 1952, pp. 111–14.

3. A Magnetically Regulated Portable Battery Charger, **R. E. D. Anderson.** *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1955 section), pp. 607–10.

4. Industrial Electronics and Control, **R. G. Kloeffler.** John Wiley & Sons, Inc., New York, N. Y., 1949, chap. XVI.

5. Silicon P-N Junction Alloy Diodes, **G. L. Pearson, B. Sawyer.** *Proceedings*, Institute of Radio Engineers, New York, N. Y., vol. 40, no. 11, Nov. 1952, p. 1348.

6. Avalanche Breakdown in Silicon, **K. G. McKay.** *Physical Review*, New York, N. Y., vol. 94, May 15, 1954, p. 94.

7. Some Applications of Semiconductor Devices in the Feedback Loop of Regulated Metallic Rectifiers, **B. H. Hamilton.** *AIEE Transactions*, vol. 73, pt. I, 1954 (Jan. 1955 section), pp. 640–45.

8. Transistors and Junction Diodes in Telephone Power Plants, **F. H. Chase, B. H. Hamilton, D. H. Smith.** *Bell System Technical Journal*, New York, N. Y., vol. 33, no. 4, July, 1954.

Discussion

J. H. Miller (Weston Electrical Instrument Corporation, Newark, N. J.): This paper is important in bringing to the attention of engineers the fact that a reasonably accurate voltage standard is now available as a circuit element and which will fill a place between the low voltage of the standard cell and the high voltage of a glow-discharge tube.

Up to this time the applications so far discussed have been largely with regard to functioning as a voltage reference from which an amplifier operates on a relatively powerful rectifier system. However, in the field of measurement there are numerous applications where a well-controlled voltage source would very considerably simplify the procedure. Resistance measurements of moderate accuracy, for example, require a comparison of the current flowing through the resistance in question to that flowing through a standard resistance with the same applied voltage, the comparison being made through the use of a balanced Wheatstone bridge or a ratio meter mechanism, of which there are several commercial varieties. The simple series ohmmeter operated by a battery is considered less accurate and generally less acceptable, principally because of battery voltage variation, which requires, in turn, that every time a measurement or a group of measurements is made that the circuit be readjusted in some manner for the existing battery voltage. But using the reverse characteristic of an alloy junction diode it now appears that voltage regulation of an entirely new order can be obtained which will allow for much higher accuracy in the over-all use of series-type ohmmeters at the expense of a limited amount of auxiliary apparatus and, it is hoped, at small additional expense.

A further important factor in the use of the reverse-current characteristic in the measurement art is the fact that the voltage values are down in the region where they match the voltages required for instrument operation; a gas tube holding a potential at 100 volts is very wasteful if a fraction of a watt is wanted, stabilized at, say, 1 volt. Further, the reverse current characteristic does not require an initial overpotential as is the case with the gas tube where it is necessary to initiate an arc before regulation commences. Thus the power source can operate, in the case of a battery supply, all the way down to the regulated potential.

At the other end of the potential spectrum, a 500-volt diode or set of series diodes would allow for simplified resistance measurement for the testing of insulation on a basis somewhat simpler than the currently used device embodying a hand-driven generator.

While perhaps not exceeding the number which will be used in the communication and power fields, it is quite probable that the various uses in the measurement field may well become quite as important.

D. H. Smith: Mr. Miller has raised some interesting aspects regarding potential uses of the silicon-alloy junction diode in the field of instrumentation. I see no reason why the diodes cannot be used as he suggests, and I share his hope that costs will be moderate.

A Survey of Magnetic Recording

S. J. BEGUN
NONMEMBER AIEE

THE understanding of magnetic recording and playback processes and the applications for magnetic recording have made considerable strides during the last 15 years. In the period of the early 1930's magnetic recorders had to operate with a recording-medium speed from 5 to 10 feet per second to secure a frequency response from 50 to 5,000 cycles per second. Today it is almost standard practice to achieve the same response characteristic with much higher signal-to-noise ratio and with considerably less distortion even though the medium moves only with a velocity of 2.5 to 3 inches per second. This represents a speed reduction of 20 to 1. The weight reduction is still greater, namely, 50 to 1 since the specific weight of plastic tape is at least 2.5 times less than that of steel tape which was used at that time. No other method of recording, whether mechanical or optical, has made such an advance in such a short period. A brief review of the history of magnetic recording indicates that 20 years ago the available magnetic-recording media with their relatively low coercive force limited the high-frequency response. In about 1936, Vicalloy tape, with more suitable magnetic properties, shifted the burden of improvements to magnetic heads. Then, coated magnetic tape and ring heads made their appearance, both intrinsically capable of a performance characteristic which has as yet not been fully exploited.

Recording Media

Of all the various magnetic-recording media which have been employed in the past, only the ones using magnetic particles imbedded in plastic materials have achieved a position of prominence. Wire and tape, whether made from a solid magnetic alloy or by plating a magnetic alloy on a nonmagnetic core, are now used only for special purposes. There are good reasons for preferring a recording medium with a thin layer of dispersed magnetic particles. Such layers can be easily applied to a wide plastic or paper base which can be cut into tapes of any desired width. These tapes have great mechanical flexibility, are relatively light, and can be conveniently handled by the user. Furthermore, they possess magnetic properties much superior to those of solid or plated tapes.

The iron-oxide particles which consti-

Paper 54-383, recommended by the AIEE Television and Aural Broadcasting Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 15, 1954; made available for printing August 16, 1954.

S. J. Begun is with the Clevite-Brush Development Company, Cleveland, Ohio.

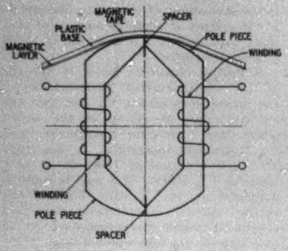


Fig. 1. **Basic structure of a magnetic ring head**

tute the magnetizable substance of the layer have an average particle size of approximately 0.4 micron. Particles rarely exceed 0.7 micron and are mostly needle-shaped with a length-to-thickness ratio of about 4 to 1. The bulk magnetic properties of these layers can be changed over a considerable range of coercivity. The most widely used tapes have a coercivity from 200 to 300 oersteds and a remanent induction of about 600 to 800 gausses. Higher coercivity does not add materially to the medium's capability of retaining short wave length signals if at the same time the remanent induction cannot be substantially increased too. Unfortunately, it is not easy to raise the remanent induction much since the coating becomes brittle when the percentage content of magnetic particles is too large. Present-day coatings contain by weight 75 to 85 per cent, and in volume 45 to 55 per cent magnetic particles. The relatively low remanent induction values are dictated by two factors, namely, the demagnetization of the material because of the nonmagnetic gaps between particles, and the reduced effective cross section since only half of the layer is occupied by particles. It would indeed be desirable to have a magnetic-recording medium with higher remanent induction and also with higher coercivity because the energy content of the reproduced signal would correspondingly increase. Much work is done these days to obtain an increase of remanent induction by more suitable shaping of particles so that they can be arranged in a pattern which will minimize the demagnetization effect. It is well to recognize, however, that for sound-recording applications magnetic media have been perfected to a point where there is now little pressure for further improvements.

This statement does not apply when magnetic-recording systems are employed for instrumentation work. So-called dropout of signals which can be traced to nonmagnetic impurities and pinholes in the coating adversely affect the use of magnetic tapes for computers and telemetery. Surface roughness, nonuniform particle distribution, and variation of coating thickness, all causing undesirable level variations, make it practically impossible to retain data with the required accuracy when the conventional intensity method commonly used for sound recording is employed. It is therefore general practice to resort to some coding method such as a frequency modulation, pulse-width-ratio modulation, or other similar schemes. Incidentally, d-c values and low-frequency phenomena can also be recorded and reproduced when carrier techniques are applied. Unfortunately, the use of many such tricks introduces new problems which are caused by speed variations of the recording medium. This will be discussed in somewhat greater detail in a later part of this paper.

It should be mentioned here that the recent introduction of DuPont Mylar polyester film as a support for a magnetic coating has greatly reduced speed variations because of its higher elastic modulus than that of acetate film which, so far, has served almost exclusively as base material for magnetic tapes. In addition, Mylar polyester film maintains great physical stability over a temperature range from −60 to 95 degrees centigrade (C). Measurements reported by members of the Reeves Soundcraft Corporation indicate that a 48 to 72 hours' exposure to 150 C caused only a 3 per cent permanent shrinkage, and a 72 hours' exposure to 125 C produces less than 1 per cent shrinkage. Since the moisture absorption of Mylar is 1/30 that of cellulose acetate, all practical problems of shrinkage and swell disappear. Recent improvements have also been made in reducing the coefficient of friction of the magnetic layer by surface polishing and by adding lubricants to the coating. This is important in counteracting flutter effects which are caused by friction-excited vibrations.

The question is frequently raised as to what extent presently available recording media might control the shortest wave length which can still be recorded and reproduced. Since, in a rough approximation, each half wave length can be assumed to equal the length of a dipole magnet, and since each particle might be considered the shortest possible magnetic dipole, particle sizes of 1/2 micron will permit recording wave lengths as short as 1 micron. This is equivalent to recording 25,000 cycles per lineal inch along the motion of the recording medium. If frequencies up to 4 megacycles have to be retained, as might be needed for storing the information content of a television program, the minimum velocity of the recording medium is approximately $13\frac{1}{2}$ feet per second.

It has been reported by members of the Radio Corporation of America that video recordings have been made with a tape speed of 30 feet per second, but more than one magnetic track has been employed. No specific information has been given regarding the highest frequency which this system is capable of recording and reproducing on anyone of the four video channels.[1,2] Present-day commercial magnetic-recording systems can resolve wave lengths of about 0.0004 inch.

Magnetic Heads

At this stage of the art, the magnetic reproducing head sets the limits. The effective gap length of the head must be smaller than the wave length which the reproducing device is expected to read. To build a practical head with a gap length of less than 1 micron is a difficult if not impossible task. The degree of required mechanical accuracy is far beyond that which can be reliably and consistently obtained. The making of heads with an extremely small gap length is further aggravated by the fact that many materials which might magnetically be suitable for pole pieces have a rather coarse grain structure. This is particularly true for the ferrites which are otherwise attractive because of their low eddy-current losses. When it comes to recording and reproducing frequencies in the megacycle range, eddy-current losses are serious. Magnetic heads made by stacking up laminations of permalloy or other high mu materials are not too practical in this frequency region. It is also necessary that the separation between magnetic layer of the recording medium and the gap be of a smaller dimension than the length of the gap.

During the last 10 years the so-called magnetic ring head has been the basis of all major head design efforts. In a magnetic ring head, the magnetic material forms an essentially closed structure with one or more nonmagnetic gaps. The magnetic-recording medium bridges one of these gaps. A schematic sketch of a balanced ring head is shown in Fig. 1. It comprises two pole pieces of equal dimension, each equipped with one winding. The effective gap length which determines the resolution power in reproduction is determined by three factors: The thickness of the spacer separating the pole pieces; the sharpness of the pole-tip corners at the point of contact with the recording medium; and the distance between the pole-piece surface and the recording medium. Fig. 2 shows the unequalized response versus frequency of a commercially available head (Brush *BK-1090*). To obtain consistently such a response curve, most careful workmanship and elaborate tooling is required. The pole faces are made flat and the corners sharp prior to inserting a close-tolerance nonmagnetic shim. The contact areas of the pole pieces are polished to improve contact conditions between the magnetic layer of the recording medium and the head. The effective gap length of this head is 3 by 10^{-4} inch and is determined by inspecting the response of the head versus wave length and by locating the first point of minimum output in the short wave length range.[3]

The playback heads are responsive only to rate of flux changes. This means that the voltage output is proportional to the frequency. In practice, this holds true only for a limited range. The reproduction of very short wave lengths brings in the gap-length effect. The reproduction of wave lengths much longer than the contact length between pole pieces and recording medium in the direction of tape motion causes a part of the recorded flux not to enter the head. Under this condition, an 18-decibel-per-octave rule applies. Magnetic heads responsive to the amplitude of flux have been built. One type employs the beam deflection of a cathode-raylike tube located between two pole pieces forming a narrow gap where they touch the medium.[4,5] Magnetic-amplifier principles are used in another approach. This is accomplished by changing the permeability of the magnetic structure of the head by means of a suitably excited high-frequency magnetizing field which does not affect the recorded signal.[6] Both these heads can read flux emanating from a stationary medium but are subject to the same limitation with regard to long and short wave lengths, as outlined before. A d-c signal, longitudinally recorded, cannot be detected by any head, since no flux leaves the surface of the medium when the internal flux density remains constant along the length of the recording track.

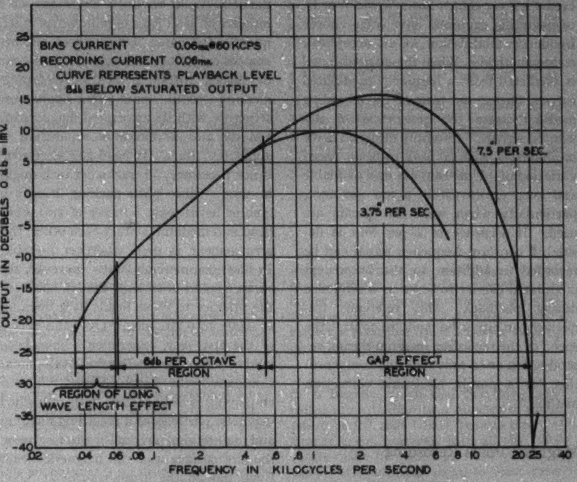


Fig. 2. Output versus frequency for BK-1090

Recording and Reproduction of Wide-Band Frequency Phenomena

Attention might now be turned to the many applications for magnetic recording. One can broadly differentiate between two major usages, namely, recording and reproduction of wide-band frequency phenomena as they occur in acoustical and other related work, and the recording and reproduction of pulses. For sound recording, only a limited degree of long-time stability is needed. It generally suffices if the relative amplitude of the various frequency components is maintained and the noise is kept small in relation to that of low-level signals and distortions do not exceed a predetermined value. Magnetic recording will meet these requirements well. While there is a tendency for short wave length signals to exhibit with time and with the number of playbacks level reductions relative to long wave length signals, these changes of one or possibly two decibels are generally not objectionable. A signal-to-noise ratio of better than 60 decibels is unique in comparison with other methods of recording. Harmonic distortions even at this excellent signal-to-noise ratio do not exceed 2 to 3 per cent in commercially available equipment. Signal transfer from one to an adjacent layer of tape, when wound on a reel, might be considered a deficiency. However, the level of the transferred signal is 49 to 57 decibels down for many commercial tapes.[7] The magnitude of the printed signal depends on the period of contact between layers, the magnetic properties of the magnetic layer and its past history, the intensity of the signal recorded thereon, the prevailing temperatures, the external magnetic field conditions, and the separation of the magnetic layers of the tape. When motion-picture film base is used as support for the magnetic layer, signal transfer is extremely small. On the other hand, the use of thinner tapes than the conventional acetate ones, such as have been proposed for Mylar polyester film, might create a serious printing problem. The application of a small erasing field prior to reproduction has been recommended since it will reduce noticeably the printed signal but will hardly affect the parent signal.[8]

In the reproduction of sound, casual defects such as caused by nonuniformity of the coating or temporary change of contact between recording head and medium are hardly observable. Experience with six magnetic sound tracks and one control track, as pioneered by Cinerama, has shown the excellent performance capabilities of magnetic recording. Cinerama's Cinemascope's demonstrations of the acoustic effect of multichannel sound systems might in due course lead to a commercial version for stereophonic reproduction in homes.

In the field of computing devices and in the area of instrumentation, magnetic recording is playing an ever-increasing part. For these applications, some method of coding is frequently required to provide the necessary degree of accuracy. In telemetery and in the recording of seismic phenomena, errors must not exceed a few per cent. Since World War II, telemetering engineers have often chosen frequency modulation as an acceptable coding method. When the original signal is extracted from a frequency-modulated carrier, flutter and wow of the recording medium cause the generation of unwanted signals. Carefully designed drive mechanisms with a peak-to-peak speed variation of 0.1 per cent are now available and keep errors within a few per cent, particularly when wide side bands are used. If accuracy of 1 per cent is desired, a constant-frequency signal can be recorded in addition to the frequency-modulated carrier to serve as a reference during playback. This reference frequency can either be made to control the speed of the recording medium during reproduction or to provide some compensation for a purely electronic correction method.[9]

Magnetic Pulse Recording

Sampling methods are becoming more accepted lately.[10] In this approach, the instantaneous amplitude of the signal is determined at precise intervals and the magnitude of each sample is expressed either as the time-duration ratio of the positive and negative pulse length of a square wave, or as a digital number. The first type of coding is identified as pulse-width-ratio modulation, and the second as pulse-code modulation. In the pulse-width-ratio modulation usually only one recording channel is used. In pulse-code modulation it becomes frequently practical to record the binary number corresponding to the amplitude of the signal on parallel channels. Eight channels will reduce the error to less than 1 per cent. In either method, it is sufficient to use only three samples for each cycle of the highest frequency component contained in the signal if the sampling is done properly.

Digital computer designers have worked extensively on pulse recording and playback techniques for the last 7 years. They frequently employ tape to store input and output information and often use drums as intermediate short-time memory devices. It is obviously desirable to pack as many pulses as possible within a given surface area of the recording medium. In accordance with present-day techniques, 200 to 250 pulses per inch can be accommodated in the direction of motion if the recording medium and head can be operated in contact, and 80 to 100 pulses if the recording speed is so high that contact cannot be maintained. The pulse density across the width of the medium is much less, particularly if accurate alignment of a number of head gaps is required as is frequently the case. Under those circumstances only 13 to 15 magnetic tracks can be placed within 1 inch of recording-medium width, since each pole-piece structure with its associated magnetizing coil dictates a minimum separation of about 0.07 inch. If head-gap alignment is not necessary, 50 tracks per inch might be used. Since pulse-packing density is greater by at least a factor of two when head and medium contact each other, this arrangement is used wherever possible. In most commercial drums, however, the surface speed is so high that it has been the practice to space heads by a distance of 0.001 to 0.002 inch. Over the period of the last year, experiments have been made by the Clevite-Brush Development Company with 1-inch diameter drums employing a rubber tire with imbedded magnetic particles as a recording medium and maintaining physical contact between the tire surface and the magnetic head. Even though the drum rotates at 3,600 rpm and, thus, has a surface speed of more than 180 inches per second, neither the head nor the tire have shown any objectionable wear and deterioration after more than three billion revolutions. Rubberlike bands with imbedded magnetic particles were developed by the Bell Telephone Laboratories for applications where long life in continuous use is required. For some years the Clevite-Brush Development Company has made magnetic bands for announcing and control systems and for delay devices.

For some applications, the use of small drums might be quite advantageous not only because more pulses can be accommodated per square inch of surface but also because the volume-to-surface ratio increases proportionally with the diameter thus making a bigger drum more wasteful in terms of space requirements. Also, the mechanical accuracy required for a small drum does not need to be as great as that for a 10- to 14-inch diameter drum which has heads slightly spaced from the periphery.

Recent Applications of Magnetic Recording

One of the great advantages in using magnetic tapes to handle input and output data of a computer lies in the fact that a 10-inch reel loaded with 1/4-inch tape and occupying a space of 16 cubic inches can retain as many bits of information as can be accommodated in a volume of approximately 3.5 cubic feet filled with typical paper cards now frequently serving as storage media. Furthermore, the random access time to magnetically recorded data is at least 16 times faster than to markings on paper cards. Tape-controlled machine tools are still in the experimental stage. Information representing the work cycle is stored on the recording medium and dictates the machine motion and on-off operations during playback.[11]

It is only recently that magnetic recording has begun to play an important part in the field of geophysical exploration work. Up to now, signals generated by explosive charges and reflected by the various boundary layers of the earth's strata are picked up by a series of geophones and are optically recorded on a suitable photographic paper. In many instances it is desirable, and often essential, to delay slightly one graphic tracing with regard to another, or to eliminate certain frequency components to obtain best results when interpreting a record by visual inspection. However, it is rarely known in advance what means ought to be applied to get a most informative record, and it becomes frequently necessary to fire a second shot so that the appropriate filters may be used between geophones and the recording equipment. This is expensive. On the other hand, if signals from geophones are first magnetically recorded, the most suitable filters and time delays can be chosen while transferring the data, as many times as necessary, from the recording medium to a visual chart record. The original magnetic record can also be preserved and used at later times when new methods of evaluation are available.

During World War II, various military agencies had found magnetic recording valuable for special applications but the total sales of all devices built in this period never exceeded a few million dollars. Even as late as 1945, magnetic recording was practically unknown to the public. There are no accurate sales figures available for 1953. Conservative estimates indicate that more than 200 million dollars' worth of magnetic sound-recording equipments and accessories alone were bought by the American public and the interest is still growing. But the real future of magnetic recording lies in the area of instrumentation and in the field of computers, business machines, and control mechanisms.

References

1. RCA DEMONSTRATES MAGNETIC TAPE RECORDING OF COLOR AND BLACK-AND-WHITE TV PICTURES. *Electrical Engineering*, vol. 73, no. 1, Jan. 1954, pp. 97–98.

2. A SYSTEM FOR RECORDING AND REPRODUCING

Television Signals, H. F. Olson, W. D. Houghton, A. R. Morgan, J. Zenel, M. Artzt, J. G. Woodward, J. T. Fischer. *RCA Review*, Princeton, N. J., vol. 15, no. 1, March 1954, pp. 3–18.

3. Studies on Magnetic Recording, W. K. Westmijze. *Phillips Research Reports*, Eindhoven, Holland, vol. 8, April, June, August, October 1953, pp. 148–57, pp. 161–83, pp. 245–69, pp. 343–66.

4. A Vacuum Tube for an Electron Beam Magnetic Reproducing Head, L. E. Leveridge. *Proceedings*, Department of Defense Symposium on Magnetic Recording, Department of the Navy, Washington, D. C., March 1954, pp. 2-1—2-13.

5. Electron-Beam Head for Magnetic Tape Playback, A. M. Skellett, L. E. Leveridge, J. W. Gratian. *Electronics*, New York, N. Y., Oct. 1953, pp. 168–71.

6. A Magnetostatic Reading Head, S. M. Rubens, A. B. Bergh. *Proceedings*, Department of Defense Symposium on Magnetic Recording. March 1954, pp. 4-1—4-15.

7. Magnetic Print-Through—Its Measurement and Reduction, Lyman J. Wiggins. *Journal*, Society of Motion Picture and Television Engineers, New York, N. Y., vol. 58, May 1952, pp. 410–14.

8. Selective Erasure of Magnetic Tape Cross-Talk, R. H. Herr, R. A. Von Behren. *Electronics*, New York, N. Y., vol. 25, August 1952, pp. 114–15.

9. F-M Recording in Guided Missiles, Albert A. Gerlach. *Electronics*, New York, N. Y., vol. 26, no. 1, Jan. 1953, pp. 108–11.

10. Radio Telemetry, M. H. Nichols, L. L. Rauch. *Review of Scientific Instruments*, New York, N. Y., vol. 22, no. 1, Jan. 1951, pp. 1–29.

11. Tape-Controlled Machines, Lawrence R. Peaslee. *Electrical Manufacturing*, New York, N. Y., vol. 52, Nov. 1953, pp. 102–08.

Discussion

C. J. LeBel (Audio Devices, Inc., New York, N.Y.): It is good, if most unusual, to see a magnetic recorder manufacturer who has kind words about magnetic recording tape. From this point of view, the most interesting part of Dr. Begun's paper is the indication that tape might be used to record up to 25,000 cycles per second per inch of tape speed. An excellent support for this figure is the fact that the Radio Corporation of America group cited seems already to be recording nearly 15,000 cycles per second per inch of speed. This is far beyond the 2,000 cycles per second per inch which has generally been regarded as a maximum.

In other matters I fear that the paper is not always on as sound ground. The author states that surface polishing reduces the coefficient of friction of a magnetic layer. This is completely contrary to our experience, for a tape which has been polished by use or by other means always shows a higher coefficient after such treatment. This is true, regardless of the manufacturer of the tape, and regardless of the binder formulation used.

Dr. Begun suggests that reduction of coefficient of friction of the magnetic layer, by adding lubricants to the coating, is a recent improvement. Audio Devices, Inc. has done this for over 6 years, which would hardly seem to make it a very recent improvement.

In his comments on dropouts the author neglects recent improvements in tape manufacturing methods. For over a year we have manufactured tape with not more than one dropout in 50,000 feet 1/4-inch width (using 0.050-inch multiple tracks for the test).

As an old disk recordist, I must challenge Dr. Begun's statement that tape's signal to noise ratio of better than 60 decibels is unique in comparison with other methods of recording. Hot-stylus recording on lacqer disks will produce with ease a signal-to-noise ratio of over 75 decibels. In this respect, it is lacquer which is unique amongst direct recording materials

S. J. Begun: I am glad that Mr. LeBel called my attention to a number of points where he felt that I have been in error. I cannot agree, however, with all the statements made by him, e.g., those related to signal dropout and signal-to-noise ratio. At least in these two instances the experience of others who have also a seasoned background in the recording field does not concur with Mr. LeBel's remarks.

By the way, the paper was not presented by "a magnetic recorder manufacturer"; I attempted to survey the magnetic-recording field in accordance with the best information available to me.

A New Standard for Direct Current

F. C. FITCHEN
ASSOCIATE MEMBER AIEE

F. P. SCHWIEG
ASSOCIATE MEMBER AIEE

W. H. TUCKER
MEMBER AIEE

Synopsis: A practical standard for direct current has been developed and is incorporated into a high-accuracy measuring instrument. Utilizing the stability of a permanent magnet, this device may be stored and used under extreme conditions of temperature, and it is relatively insensitive to variation in operating conditions.

A STUDY of the measurement and control of process variables led to the conclusion that a significant contribution could be made to this art, particularly in the design of a measurement system susceptible of use over wider ranges of temperature than those in use today. The ruggedness of critical components of such a system was also a consideration of extreme importance. A new measurement system for potentiometer measurement has been described in another paper.[1] An element of great importance in this system is a standard for direct current employing a well-stabilized permanent magnet.[2]

The function of this component is to relate the magnitude of direct current and the magnitude of physical angle of a shaft carrying a stable permanent magnet. An inequality results in an electrical actuating signal which, in the aforementioned measurement system, is used in an amplifier-motor combination to correct the angle until the actuating signal is reduced to zero. In the balanced condition the angle is a precise measurement of the current. This arrangement is shown schematically in Fig. 1.

Since the magnetic standard may be used in practice to effect a relationship of current with angle or of current with current, other interesting applications, such as in a constant current device, in a precise telemeter of angle, and in a current comparator or repeater have been made.

Objective

The objective for such a magnetic standard included the following:

1. The precise relating of direct current with shaft angle, including a proportional relationship.

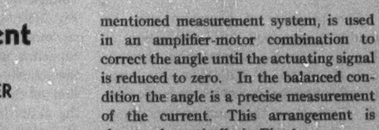


Fig. 1. Block diagram of d-c measuring system employing the magnetic standard

---

Paper 54-354, recommended by the AIEE Instruments and Measurements Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted July 12, 1954; made available for printing August 8, 1954.

F. C. Fitchen, F. P. Schwieg and W. H. Tucker are with the General Electric Company, West Lynn, Mass.

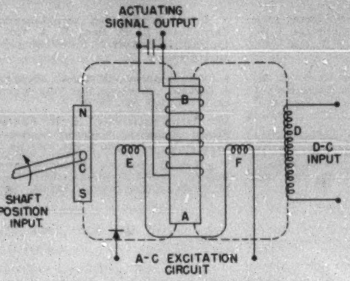


Fig. 2 (above). Rudimentary magnetic standard

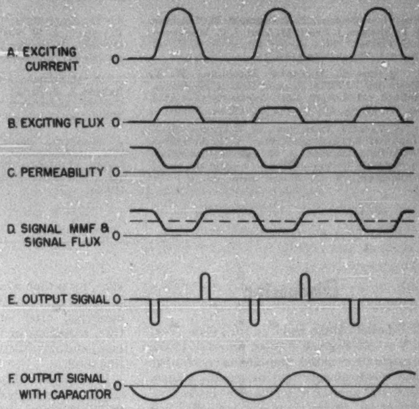


Fig. 3 (right). Wave forms of electric and magnetic quantities in the standard

2. Stability of calibration of 0.02 per cent for a period of several years.

3. Ability to withstand temperature extremes of −25 and +70 degrees centigrade without permanent change in calibration, storage, and operation.

4. Suitable physical dimensions for inclusion in a process control instrument of conventional size.

5. Production of a usable signal for servomotor operation for a condition of unbalance of 0.02 per cent of full scale value.

6. Economical production.

Principle of Operation

The principle of operation may be readily understood by examination of Fig. 2. The d-c input acting through coil *D* creates a magnetomotive force (mmf) which is applied to magnetic core *A*. To balance this mmf another input is applied by means of magnet *C*, which is rotatable upon a shaft whose axis lies in the plane of the paper. The shaft position may therefore be varied until the net mmf acting on the core is zero. For any condition of unbalance, there is a net flux in the core.

To obtain an a-c actuating signal in the winding *B* upon the core, it is necessary to cause a cyclic variation of the permeability of the core. This is done in Fig. 2 by cross-magnetization of the core using coils *E* and *F*. If these coils are excited with a sufficiently high level of half-wave rectified current, their mmf will saturate the core once during each cycle, thus lowering its permeability once per cycle. The net mmf resulting by subtracting that of the direct current from that of the magnet then produces a cyclically varying flux in coil *B*, which produces an electromotive force at the terminals of this coil which is of the same frequency as the excitation supply. Fig. 3 illustrates the wave forms of the various electric and magnetic quantities involved. Fig. 3(A) illustrates the variation of excitation current in coils *E* and *F*. Fig. 3(B) shows the exciting flux in core *A*, and Fig. 3(C) shows the variation of permeability in the core as a result of saturation. Fig. 3(D) illustrates a condition of unbalance as shown by the unvarying signal mmf and the cyclically varying signal flux which follows the variation in permeability of the core. This variation of flux induces a voltage in winding *B*, as illustrated in Fig. 3(E). The actuating signal may be smoothed by application of a suitable capacitor across the coil terminals. If the polarity of the signal mmf is reversed, that of the signal flux also reverses and the actuating signal voltage changes in phase 180 degrees from the voltage illustrated. In this way, the device creates a signal voltage which is a measure of the sense as well as of the magnitude of the unbalance.

Construction

For numerous practical reasons, the device in use, illustrated in Fig. 4, differs materially in construction from the illustrative sketch of Fig. 2. C. W. LaPierre[3] suggested the possibility of combining the input coil and the actuating signal coil; see the Appendix for details. In practice, this coil is divided into two equal portions which are wound upon the outer legs of a silicon-steel punching, the general shape of which is shown in Fig. 5. These coils are so connected that the fluxes produced in the legs are as illustrated and a partial return path for flux is provided by a can of magnetic material which also serves to shield the device from external fields. In Fig. 5, the magnet is shown in the position where it exerts no useful mmf effect upon the core. When the magnet is turned from this position, a part of its flux can be caused to follow the core, as illustrated by the dotted arrows in Fig. 5.

Thus, for each value of current within the range of the device, there is a value of angle of the magnetic which will produce zero flux in the outer legs of the core. The recognition of unbalance is achieved by adding a pair of windings connected as shown in Fig. 6 and excited through a half-wave rectifier. These windings create the exciting flux, whose entire path is in the core, which induces in the input-output windings electromotive forces which cancel one another because of the relative polarities of connections shown. The wave forms of Fig. 3, when considered in terms of the outer legs of the punching, apply as they did to Fig. 2.

A study was made of the effect of variation of permeability of steel with time and the effect of clamping silicon-steel stacks. One basic conclusion reached was that best performance could be obtained by use of a single lamination in which stress could be minimized by clamping it to the supporting frame at one point only. The spools are supported by the frame and do not touch the lamination.

It has already been noted that the coils of the system encircle the outer legs of the lamination and that a partial return path for flux from the d-c coil is provided by the shield can. This return path effectively prevents flux from the d-c coil from cross-magnetizing the magnet. Significant for operating the device are the keyhole-shaped slots which lie in the core poles. These cause a continued washing of the poles by the excitation flux and re-

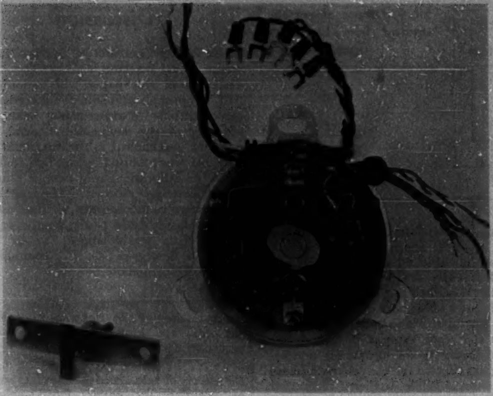


Fig. 4. Interior view of standard showing magnet, lamination, and coils

duce to an insignificant amount any tendency to residual magnetism that might be expected because of the proximity of the poles to the permanent magnet.

The addition of the rectifier in the excitation circuit[2] has proved to be an important contribution since it created an actuating voltage of the frequency of excitation and, at the same time, rendered this voltage relatively independent of variations in magnitude of the excitation voltage. A reinspection of Fig. 3 will suggest that, with a-c excitation and without the rectifier, the core flux would rise twice per cycle to saturation. Therefore, the permeability would fluctuate at second-harmonic rate and produce a second-harmonic actuating signal. The relative difficulties of amplification and servomotor operation using a second-harmonic signal are widely known, and the cost of the instrument in which the standard is used would have been excessive had it been necessary to provide a second-harmonic reference voltage for the servomotor.

Because of the impracticability of precise duplication of the linear characteristics, a silicon-steel adjuster, called the trim, was added. This may be seen disassembled in Fig. 4. When in use, it is brought into relationship with the magnet, as shown in Fig. 7. It has the property of inserting an adjustable bow of ±0.15 per cent maximum in the characteristic, as illustrated in Fig. 8. The magnet and core shape were so chosen that the average characteristic with the trim at the center of its adjustment is truly linear. The trim then is used to rectify the characteristic where combinations of parts assembled at random show a resultant characteristic lying beyond 0.05 per cent from the average.

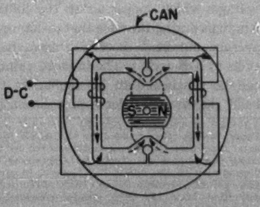


Fig. 5. Schematic of magnetic circuit and input windings

One significant quality of a device of this precision is the effect upon calibration of the ambient temperature. This arises chiefly from the change of strength of the magnet and was found to be for this device approximately −0.018 per cent per degree centigrade. To reduce this error, some experimental work was done with the Curie alloy of iron and nickel which has been used successfully in other applications.[4] It was determined, however, that a practical compensation could be achieved by electrical means, and a high-resistance shunt of negligible temperature coefficient was connected across the d-c coils. With respect to this, it must be noted that the device demonstrates a temperature memory effect on the order of 0.02 per cent.

One characteristic studied by researchers in a similar investigation[4] is the quadrature component signal which appears as a broad null when investigated qualitatively. This signal has a tendency, when large, to saturate the input of the amplifier. When used with a direct-reading null detector, it makes the determination of the balance point less well defined. A method of reducing quadrature signal was investigated and as a result a tertiary winding of few turns was added to each outer leg of the core. In production, one of these windings is closed experimentally through an external resistor to reduce the quadrature voltage in those units where it is excessive. The quadrature appears to be linked to those constructions in which the magnetic configuration is not completely symmetrical.

Performance Characteristics

The new standard has an angle-to-current relationship for 60 degrees of angle and 0 to 5 milliamperes of current with a maximum deviation from linearity, occurring at half-scale, of 0.05 per cent. It has an error not greater than ±0.5 per cent for changes in excitation voltage of ±10 per cent. The ambient temperature error is less than 0.05 per cent at temperature extremes of −18 to 50 degrees centigrade.

Installed in an instrument case, it has an error caused by stray field of less than

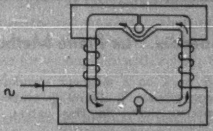


Fig. 6. Lamination and exciting winding

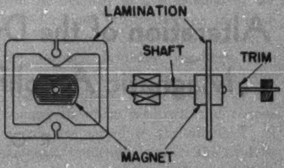


Fig. 7. Physical relationship of trim to magnet and lamination

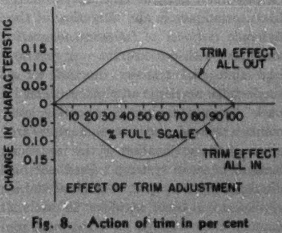


Fig. 8. Action of trim in per cent

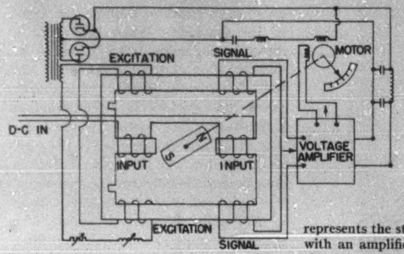


Fig. 9. Diagram of magnetic standard from La Pierre patent

0.04 per cent for a 5-gauss field of 60-cycle frequency and for a 5-gauss d-c excited field. Life tests have demonstrated that the device has a stability which is of the order of accuracy of measurement (0.02 per cent) for periods in excess of 2 years, even when subjected during this period to weekly temperature cycles of −20 to 50 degrees centigrade.

Conclusion

The new standard is believed susceptible of use in many applications requiring a rugged, compact comparator insensitive to many interfering effects, including those of temperature, shock, gravity, vibration, etc.

Appendix. La Pierre Method

Basic work on a magnetic standard is given in reference 3. Fig. 9 illustrates the construction shown in this patent and represents the standard used in conjunction with an amplifier and a 2-phase motor for the measurement of direct current. It will be noted that the core configuration is similar to that described in this paper but that the d-c input and signal windings are separate and the input windings are placed upon poles on the core. Also of interest is the indicated necessity of providing double-frequency excitation for the reference phase of the motor. While LaPierre illustrated reduced sections in the core to provide saturation, his patent covered core forms of other natures so designed as to provide appropriate saturation. He also pointed out the possibility of combining the functions of the input and actuating signal coils.

References

1. Advances in Null-Type Recorder-Controller Design, **I. F. Kinnard, M. A. Princi, A. Hansen, Jr.** *ASME Paper No. 54-IRD-12*, American Society of Mechanical Engineers, New York, N. Y., 1954.

2. Saturable Magnetic Core Control Device, **F. P. Schwieg.** U. S. Patent *2,674,705*, April 6, 1954.

3. Direct-Current Indication, **Cramer W. LaPierre.** U. S. Patent *2,053,154*, Sept. 1, 1936.

4. Preliminary Development of a Magnetor Current Standard, **E. P. Felch, J. L. Potter.** *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 524–31.

Discussion

L. E. Rinker (Leupold & Stevens Instruments, Inc., Portland, Ore.): The device described in this paper will undoubtedly find many applications in the recorder and servocontroller measurements field. The accuracy of the device will depend upon precision mechanical fabrication, and it is also somewhat limited by the fact that it is linear in calibration only over a range of approximately 60 degrees. The fact that it is necessary to incorporate a "trim" adjustment makes it subject to human error and, therefore, it would appear to be more properly described as a "comparator" instead of a "standard."

F. C. Fitchen, F. P. Schwieg, and **W. H. Tucker:** Mr. Rinker's comments are appreciated. It was not possible in the body of the paper to discuss at length the effects of tolerances of manufacture and of assembly of parts. Experience has confirmed predictions that the only critical procedures are care in handling the lamination after annealing, and in handling the magnet after charging.

Increased scale length for readability is possible without loss of accuracy through the use of antibacklash gearing.

The authors do not feel that the existence of the trim adjustment should deny this device the title of standard, since all standards are subject to human error. For example, compare with the American Standards Association's definition for resistance standard which states "A resistor which is adjusted with high accuracy to a specified value. . ."[1]

Reference

1. American Standards Definitions of Electrical Terms. *ASA C42-1941*, American Standards Association, New York, N. Y., 1941.

Alteration of the Dynamic Response of Magnetic Amplifiers by Feedback

R. O. DECKER
ASSOCIATE MEMBER AIEE

THE purpose of this paper is to explain and illustrate the use of feedback techniques in the alteration of the dynamic response of full-wave magnetic amplifiers with d-c outputs. A generalized sinusoidal transfer function for a single-stage magnetic amplifier with feedback through frequency sensitive elements is derived, and several examples are given of the application of the feedback method in servo systems. Feedback, as discussed in this paper, is external and is applied to a basic amplifier in such a manner as to change its frequency response characteristics. The self-saturating amplifier, in particular, is considered to be a basic amplifier to which negative or positive feedback may be applied, although the analysis is also valid for other types of magnetic amplifiers.

The transfer functions developed in the paper are sinusoidal. They describe the steady-state response of the amplifier to a sinusoidal control signal.[1] The frequency response approach is used because the analysis and design of servo systems based on the frequency analysis method can be carried out in a straightforward manner when the transfer functions of the system elements are known.[2]

Transfer Function for Magnetic Amplifier Without Feedback

A magnetic amplifier of the full-wave self-saturating type has a delay similar to that of an inductive circuit if the residual delay of the amplifier is neglected. The inductive delay is caused by the induced voltages which appear across the control windings during the time that a change in average output is occurring from half-cycle to half-cycle. The induced voltage can be considered to be similar to the back electromotive force of a linear in-

---

Paper 54-376, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted May 18, 1954; made available for printing August 13, 1954.

R. O. Decker is with the Westinghouse Electric Corporation, East Pittsburgh, Pa.

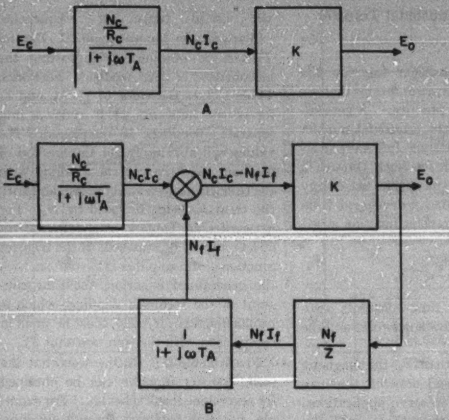

E_c
$\frac{N_c}{R_c} / 1+j\omega T_A$
N_cI_c
K
E_o
A
E_c
$\frac{N_c}{R_c} / 1+j\omega T_A$
N_cI_c
$N_cI_c - N_fI_f$
K
E_o
N_fI_f
$\frac{1}{1+j\omega T_A}$
N_fI_f
$\frac{N_f}{Z}$
B

Fig. 1. Block diagram of magnetic amplifier

A—Without feedback
B—With feedback into separate windings

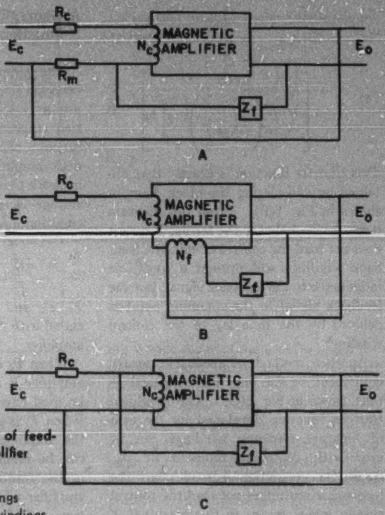


Fig. 2 (right). Three methods of feedback around magnetic amplifier

A—Resistance coupling
B—Separate feedback windings
C—Feedback into control windings

ductive circuit with an apparent inductance L_c proportional to N_c^2 where N_c is the number of control turns.[3] The time constant of the amplifier is then

$$T_A = \frac{L_c}{R_c} \qquad (1)$$

where L_c is proportional to N_c^2 and R_c is the control circuit resistance.

The gain of the magnetic amplifier can be specified in several ways, but it is convenient to use a gain factor K which relates the change in average output voltage to the change in average control ampere turns. This factor K is a constant over a considerable range of output voltage for a given amplifier and a given load although, in general, it depends on core and rectifier characteristics and load impedance. The sinusoidal transfer function of the amplifier can be written as follows

$$\frac{E_0}{E_c}(j\omega) = \frac{K\frac{N_c}{R_c}}{1+j\omega T_A} \qquad (2)$$

where K is a gain constant in volts per ampere-turn; E_0 is the average output voltage; E_c is the average control voltage; and T_A is the time constant of the amplifier proportional to (N_c^2/R_c). The block diagram of the amplifier is shown in Fig. 1(A).

The residual delay, which does not appear in equation 2, is a fixed delay which occurs because the flux level of a core is set during 1 half-cycle of the supply frequency and a change in output voltage occurs on a second half-cycle. The effect of this additional delay will be to increase the phase shift of the amplifier; thus the phase shift calculated from equation 2 will be in error. However, this error is small when T_A is several times the period of 1 cycle of the supply frequency, and in the analysis presented in this paper the amplifier is treated as a device with a single time constant.

Transfer Function for Magnetic Amplifier With Feedback

To achieve drift and gain stability, d-c negative feedback is widely used in self-saturating magnetic amplifier applications, but very little has been published concerning the use of positive or negative feedback through frequency sensitive elements. The dynamic response of the amplifier can be altered by this technique to fit a particular servo application. Steady-state feedback networks can be combined with capacitors or inductors to produce amplifier attenuation and phase characteristics that will reshape the open-loop frequency locus of a servo system to obtain dynamic stability. The magnetic amplifier is well suited to this application because the feedback quantity may be resistance-coupled into the control circuit, or feedback into isolated windings may be employed.

One type of feedback is accomplished by feeding back a portion of the output voltage directly into the control circuit, with resistance coupling, as diagrammed in Fig. 2(A). If the control circuit does not load the output circuit, the transfer function can be expressed in the usual manner for an amplifier with voltage feedback.[4]

However, this method of feedback is not as flexible as the method of feeding back into separate windings, as shown in Fig. 2(B), and the major portion of this paper will be devoted to a discussion of the considerations involved when the feedback quantity is coupled into the amplifier through feedback windings. The first step to be taken in discussing this type of feedback is the analysis of the manner in which feedback and control signals are mixed in the amplifier. A convenient starting point for this analysis is to assume that the control and feedback windings can be represented as linear inductances, proportional to N_c^2 and N_f^2 respectively, and these linear inductances can be considered to be mutually coupled in the usual manner.

In the Appendix an analysis is carried out for the case of an impedance Z placed in series with the feedback windings. The following transfer function is derived for the amplifier with the imposed condition that the feedback circuit does not load the control circuit. Stated another way, (N_c^2/R_c) must be several times larger than (N_f^2/Z) over the frequency range of interest.

$$\frac{E_0}{E_c}(j\omega)=\frac{N_c}{R_c}\left[\frac{1}{1+j\omega\frac{L_c}{R_c}}\right]\times\left[\frac{K}{1+K\left(\frac{1}{1+j\omega\frac{L_c}{R_c}}\right)\left(\frac{N_f}{Z}\right)}\right] \quad (3)$$

This transfer function is exactly that obtained by considering the block diagram shown in Fig. 1(B). From this diagram it can be seen that if the feedback circuit does not load the control circuit, the feedback windings will present a negligible inductance to the feedback signal, but the feedback signal to the summing point is delayed by the time lag of the control winding.

Another method of feedback, nearly equivalent to the one just discussed, is that shown in Fig. 2(C). The simplified transfer function developed for the case of separate feedback windings can be used with N_f being replaced by N_c. Again the restriction is made that the feedback circuit does not load the control circuit, which means in this case that Z must be several times larger than R_c. In addition, the restriction must be imposed that the impedance of the control circuit external to the control winding is large enough that the feedback ampere turns are equal to (E_0N_c/Z).

A Generalized Sinusoidal Transfer Function

The simplified transfer function has been developed with assumptions that are usually justified in practice. A generalized equation for the sinusoidal transfer function of a single-stage full-wave magnetic amplifier with feedback through a network that has no more than one energy storage element can be developed from equation 3. The transfer function is

$$\frac{E_0}{E_c}(j\omega)=K^1\frac{1+j\omega T_1}{1+j\omega T_2-\omega^2T_1T_3} \quad (4)$$

T_1, T_2, and T_3 are time constants associated with the feedback networks and the amplifier.

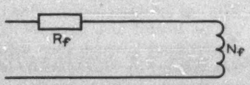
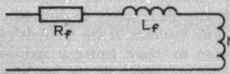
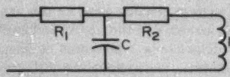
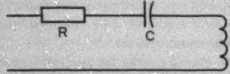
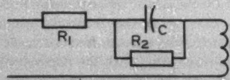
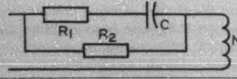
When T_1 is less than T_2, the magnetic amplifier is a lag-lead network that can be used effectively in servo applications. When T_1 is greater than $(T_2/1-\omega^2T_1T_3)$, the amplifier is a lead-lag network. It can be readily seen that the usefulness of the feedback method in making the amplifier a lead network is limited by the term $\omega^2T_1T_3$. When $\omega^2T_1T_3$ equals 1, the denominator of equation 4 has a phase angle of 90 degrees, and the net phase angle is negative even though the time constant T_1 is very large. Furthermore, increasing T_1 to obtain more phase lead increases the term $\omega^2T_1T_3$ and shortens the useful range of frequencies.

Variation in the values of T_1, T_2, and T_3 can be obtained by adjusting the parameters of the feedback networks. Table I gives the value of T_1, T_2, and T_3 for several networks when feedback into separate windings is employed. The values will also apply for the method of feedback into the control winding, with the restrictions previously noted, and with the term N_f being replaced by N_c. For network C of Table I, an additional restriction must be imposed. If the transfer function of the amplifier is to conform with the generalized equation, the time constant of the feedback winding, which is proportional to (N_f^2/R_f), must be small in comparison with the time constant T_1.

From Table I it can be seen that the same transfer function can be obtained for several feedback schemes. For example, negative feedback through an integrating network such as network B or C will give the amplifier a phase lead characteristic, with the limitations noted above. Positive feedback through a differentiating network such as network D will have the same effect, although it is generally preferable to use negative feedback because the possibility exists of converting the amplifier into an oscillator with the positive feedback method.

The advantages of the feedback method

Table I. Feedback Schemes

	Feedback Network	Type of Feedback	T_1	T_2	T_3	K'
A		Negative	0	$\frac{T_A}{1+K(N_f/R_f)}$		$\frac{KN_c/R_c}{1+(KN_f/R_f)}$
		Positive		$\frac{T_A}{1-K(N_f/R_f)}$		$\frac{KN_c/R_c}{1-(KN_f/R_f)}$
B		Negative	$\frac{L}{R_f}$	$\frac{T_1+T_A}{1+K(N_f/R_f)}$	$\frac{T_A}{1+(KN_f/R_f)}$	$\frac{KN_c/R_c}{1+(KN_f/R_f)}$
		Positive		$\frac{T_1+T_A}{1-K(N_f/R_f)}$	$\frac{T_A}{1-(KN_f/R_f)}$	$\frac{KN_c/R_c}{1-(KN_f/R_f)}$
C		Negative	$\left(\frac{R_1R_2}{R_1+R_2}\right)C$	$\frac{T_1T_A}{1+KN_f/R_f}$	$\frac{T_A}{1+(KN_f/R_f)}$	$\frac{KN_c/R_c}{1+(KN_f/R_f)}$
		Positive		$\frac{T_1+T_A}{1-K(N_f/R_f)}$	$\frac{T_A}{1-(KN_f/R_f)}$	$\frac{KN_c/R_c}{1-(KN_f/R_f)}$
D		Negative	RC	$T_1+T_A+KN_fC$	T_A	KN_c/R_c
		Positive		$T_1+T_A-KN_fC$	T_A	KN_c/R_c
E		Negative	$\left(\frac{R_1R_2}{R_1+R_2}\right)C$	$\frac{T_1+T_A+KN_fC(R_2/R_1+R_2)}{1+(KN_f/R_1+R_2)}$	$\frac{T_A}{1+[KN_f/(R_1+R_2)]}$	$\frac{KN_c/R_c}{1+[KN_f/(R_1+R_2)]}$
		Positive		$\frac{T_1+T_A-KN_fC(R_2/R_1+R_2)}{1-(KN_f/R_1+R_2)}$	$\frac{T_A}{1-[KN_f/(R_1+R_2)]}$	$\frac{KN_c/R_c}{1-[KN_f/(R_1+R_2)]}$
F		Negative	R_1C	$\frac{T_1+T_A+KN_fC(R_1+R_2/R_2)}{1+(KN_f/R_2)}$	$\frac{T_A}{1+(KN_f/R_2)}$	$\frac{KN_c/R_c}{1+(KN_f/R_2)}$
		Positive		$\frac{T_1+T_A-KN_fC(R_1+R_2/R_2)}{1-(KN_f/R_2)}$	$\frac{T_A}{1-(KN_f/R_2)}$	$\frac{KN_c/R_c}{1-(KN_f/R_2)}$

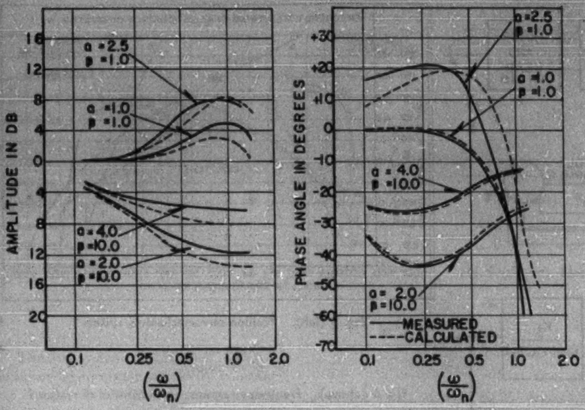


Fig. 3. Amplitude and phase plots for 60-cycle amplifier

in obtaining system compensation is readily shown by an examination of the terms in Table I. In obtaining phase lead by network *C*, the lead time constant T_1 is the product of *R* and *C*. *R* can be very large since in general *K* is large and within limits N_f may be made large. In networks *D* and *E* the major lag time constant is KN_fC; and here again the product KN_f, which has the dimension is resistance, may be made very large, with the result that *C* may be made small. For example, a typical value for the product KN_f is 100,000 ohms. This contrasts with typical series resistance levels of 1,000 to 10,000 ohms in the control circuits of magnetic amplifiers. Another advantage of the feedback method is that filtering action is readily obtained. When network *C* is used to obtain phase lead, the magnetic amplifier and the integrating network serve to filter out much of the noise which is usually a by-product of series differentiating circuits. Also, the possibility of combining d-c negative feedback with frequency compensation is an attractive scheme because of the stabilizing effect on gain and drift. The core and rectifier materials used in magnetic amplifiers are temperature sensitive; and if stability over a wide temperature range is desired, d-c negative feedback is needed.

Experimentation

To check the validity of the generalized transfer function, a 60-cycle and a 400-cycle push-pull magnetic amplifier with reversible polarity d-c output were used for testing purposes. The first step in determining the characteristics of interest was to measure the volts per ampere turn gain of each amplifier. Next, the time constant was measured by applying a small sinusoidal test signal to the control circuit and the variations in the output voltage measured by placing a nonloading, calibrated resistance-capacitance filter on the output with a time constant such that sinusoidal variations in the envelope of the average voltage could be measured.

To facilitate comparison of experimental and theoretical data, the generalized transfer function was put in the following normalized form

$$\frac{E_0}{E_c}(j\omega)=\frac{K^1\left[1+j\left(\frac{\omega}{\omega_n}\right)\alpha\right]}{\left[1+j\left(\frac{\omega}{\omega_n}\right)\beta-\left(\frac{\omega}{\omega_n}\right)^2\right]} \qquad (5)$$

Here

$$\omega_n=\frac{1}{\sqrt{T_1T_3}}$$

$$\alpha=\sqrt{\frac{T_1}{T_3}}$$

$$\beta=\frac{T_2}{\sqrt{T_1T_3}}$$

For a wide range of values of α and β the amplitude and phase angle of the normalized equation versus the normalized frequency ratio (ω/ω_n) were computed by the use of a digital computer, and curves were plotted of the amplitude and phase angle of the normalized equation. For a number of curves, the frequency-sensitive portion of the normalized equation was checked by experimental testing on the 60- and 400-cycle push-pull amplifiers. The basic procedure used in the experimental testing was to select a network and adjust its parameters to obtain certain values of α and β. The values of *K* and N_f for each amplifier were fixed, although T_A was varied by using several different values of R_c.

Experimental and calculated curves for various values of α and β are plotted in Figs. 3 and 4. It can be seen that good correlation has been obtained between the theoretical expressions and experi-

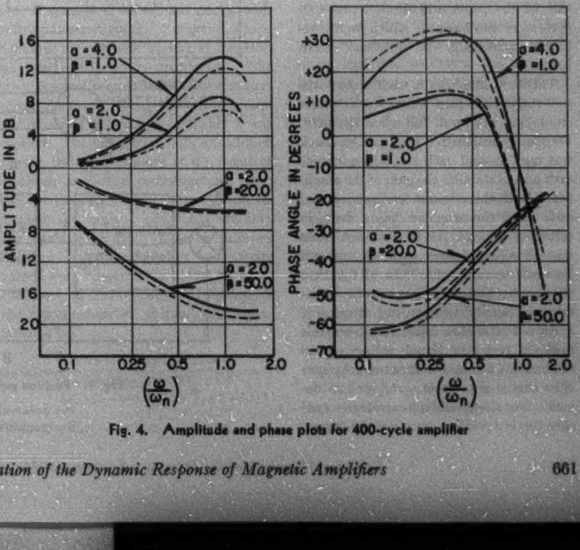
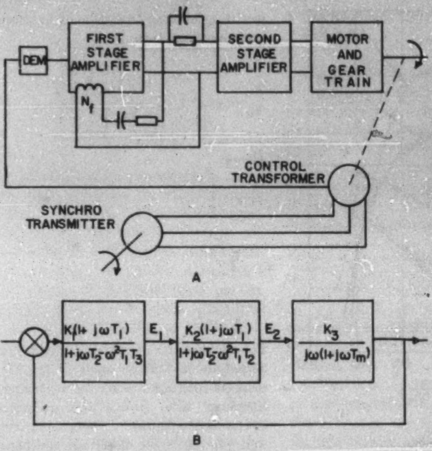


Fig. 4. Amplitude and phase plots for 400-cycle amplifier


DEM
FIRST STAGE AMPLIFIER
SECOND STAGE AMPLIFIER
MOTOR AND GEAR TRAIN
N_f
CONTROL TRANSFORMER
SYNCHRO TRANSMITTER
A
$\frac{K_1(1+j\omega T_1)}{1+j\omega T_2-\omega^2 T_1 T_3}$
E_1
$\frac{K_2(1+j\omega T_1)}{1+j\omega T_2-\omega^2 T_1 T_2}$
E_2
$\frac{K_3}{j\omega(1+j\omega T_m)}$
B

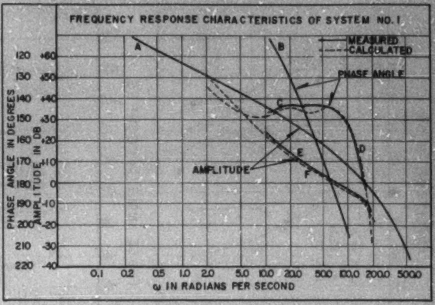


Fig. 5 (left). Position servomechanism, system 1

A—Schematic diagram
B—Block diagram

Fig. 6 (above). Frequency response characteristics of system 1

A, B—Amplitude and phase curves of uncompensated system
C, D, E, F—Amplitude and phase curves of compensated system

mental data. The small differences are due to second-order effects which have been neglected in the analysis, such as the residual delay of the amplifier and the loading effect of the feedback windings. In particular, it has been found that for extreme values of *R* and *C* in network *D* of Table I, the zero-frequency gain of the amplifier is decreased by adding the feedback network.

System Application

The feedback method has proven to be of great value in providing compensation in a wide variety of regulating and servomechanism systems, and this portion of the paper will discuss two 400-cycle servomechanisms which were stabilized principally by feedback techniques.

The servo amplifier in each system was a 2-stage high-gain magnetic amplifier employing push-pull full-wave circuitry. Frequency insensitive negative feedback was used around each stage to minimize drift and to stabilize the gain of the amplifier. With three-to-one feedback around each stage the amplifier has a drift, in terms of control current, of not more than 5 microamperes over an ambient temperature range of −65 to +71 degrees centigrade, and with 10-per-cent variations in supply voltage and frequency. With a 10,000-ohm input resistance this drift, in terms of signal voltage, is 50 millivolts. For a 1-volt-per-degree synchro this is a position error of 0.05 degree. For a 400-millivolt-per-degree synchro this is a position error of 0.125 degree.

The first step in stabilizing the systems was to measure the frequency response characteristics of each amplifier stage and the motor-gear train-load combination. Frequency response plots of the combined transfer function were then made, and stabilizing networks were chosen after a few trial calculations, which were made using the equations developed in this paper. Frequency-response plots of the compensated systems were then made to predict the stability of the closed loops. The chosen networks were inserted in the system, and open-loop measurements made to check the theoretical transfer function.

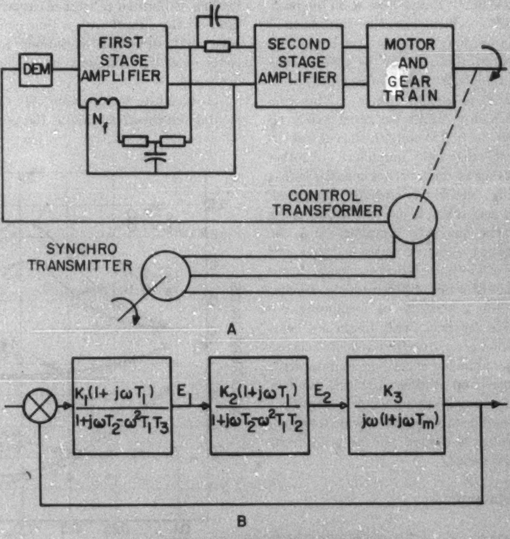
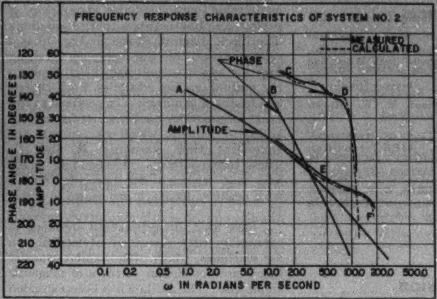

DEM
FIRST STAGE AMPLIFIER
SECOND STAGE AMPLIFIER
MOTOR AND GEAR TRAIN
N_f
CONTROL TRANSFORMER
SYNCHRO TRANSMITTER
A
$\frac{K_1(1+j\omega T_1)}{1+j\omega T_2-\omega^2 T_1 T_3}$
E_1
$\frac{K_2(1+j\omega T_1)}{1+j\omega T_2-\omega^2 T_1 T_2}$
E_2
$\frac{K_3}{j\omega(1+j\omega T_m)}$
B

Fig. 7. Position servomechanism, system 2

A—Schematic diagram
B—Block diagram



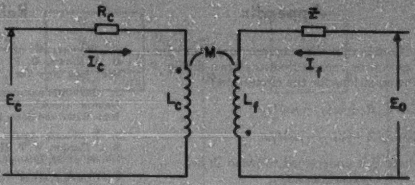


Fig. 8 (left). Frequency response characteristics of system 2

A, B—Amplitude and phase curves of uncompensated system
C, D, E, F—Amplitude and phase curves of compensated system

Fig. 9 (above). Idealized representation of magnetic amplifier winding

$$L_c = aN_c^2$$
$$L_f = aN_f^2$$
$$M = \sqrt{L_cL_f} = aN_cN_f$$

The first system, diagrammed in Fig. 5, was set up to obtain a velocity constant of 700 seconds^{-1}. The velocity constant, which is high for a magnetic amplifier system, was obtained with an amplifier voltage gain of 200 and a gear ratio of 50 between the motor and control transformer. The transfer function of the uncompensated system was determined experimentally to be

$$G_1G_2G_3 = \frac{700}{(1+j\omega 0.005)(1+j\omega 0.003)(1+j\omega 0.03)(j\omega)}$$

This transfer function is drawn as a solid line in Fig. 6. The uncompensated system has a crossover frequency of 125 radians per second and a phase margin of −45 degrees, indicating an extremely unstable condition if the loop is closed with no attempt at compensation.

The high crossover frequency, 20 cycles per second, made compensation of this system especially difficult, and the system locus was reshaped to give a lower crossover frequency. This was done by making the first-stage amplifier a lag-lead network by using negative feedback through network *D* of Table I. Additional phase lead was obtained by a series network between stages, which changes the transfer function of the second-stage amplifier from a single-delay transfer function to an expression similar in form to that of the generalized transfer function. Network parameter values were chosen after a few trial calculations, and the transfer function of the compensated system was calculated to be

$$G_1G_2G_3 = \frac{700(1+j\omega 0.056)(1+j\omega 0.01)}{(1+j\omega 0.5-\omega^2 0.0003)(1+j\omega 0.003-\omega^2 0.00003)(j\omega)(1+j\omega 0.003)}$$

This function is plotted in Fig. 6 as a dashed line. The calculated crossover frequency is 63 radians per second; the calculated phase margin is 35 degrees; and the calculated gain margin is 8 decibels (db). Experimental measurements showed a crossover frequency of 70 radians per second, a phase margin of 35 degrees, and a gain margin of 6 db.

The second system, diagrammed in Fig. 7, uses a larger motor. The system was set up to have a velocity constant of 150 seconds^{-1}. This velocity constant was obtained with an amplifier voltage gain of 125 and a gear ratio of 50 between the motor and control transformer. The transfer function of the uncompensated system was determined experimentally to be

$$G_1G_2G_3 = \frac{150}{(1+j\omega 0.005)(1+j\omega 0.003)(1+j\omega 0.1)(j\omega)}$$

This transfer function is plotted in Fig. 8. The uncompensated system has a crossover frequency of 35 radians per second, with a phase margin of −2 degrees; which indicates an unstable system if the loop is closed with no compensation. The low crossover frequency of the system suggested the use of a lead network to obtain stability and network *C* of Table I was chosen. The value of R_1+R_2 was chosen to give the same amount of d-c negative feedback as before compensation; thus the low-frequency gain of the system remained the same. Additional stability was again obtained by a series network between stages. Network parameter values were chosen after a few trial calculations, and the transfer function of the uncompensated system was calculated to be

$$G_1G_2G_3 = \frac{150(1+j\omega 0.025)(1+j\omega 0.004)}{(1+j\omega 0.004-\omega^2 0.00004)(1+j\omega 0.005-\omega^2 0.00002)(j\omega)(1+j\omega 0.1)}$$

This transfer function is drawn as a dashed line in Fig. 8. The calculated crossover frequency is 50 radians per second; the calculated phase margin is 44 degrees; and the calculated gain margin is 5 db. Experimental measurements on the compensated system show a crossover frequency of 54 radians per second, a phase margin of 40 degrees, and a gain margin of 5 db; thus the system is effectively damped.

These systems are but two of many that could be cited. They were chosen as examples to show two different methods of stabilizing. In the first system the crossover frequency was too high to allow phase lead networks to be used effectively; therefore, the system locus was reshaped to yield a lower crossover frequency. The additional phase lead necessary to make the system well stabilized was then obtained by a series-compensating network between stages. In the second example the crossover frequency of the uncompensated system was low; thus it was possible to obtain phase lead and to extend the crossover frequency by a combination of feedback around the first stage and by series compensation.

Conclusions

A generalized transfer function for a magnetic amplifier with feedback has been derived, and the theoretical expression has been verified by experimental testing. It has been shown that, although the time constant of the amplifier is a limiting factor on the performance, some of the limitations can be overcome by the proper use of feedback. Several examples have been given which should serve as a guide in the application of feedback techniques to magnetic amplifiers in servo systems.

Appendix

Assume that the magnetic-amplifier windings can be represented as shown in Fig. 9. The equations for the circuit are

$$E_c = I_c(R_c + j\omega L_c) - j\omega M I_f$$

$$E_0 = I_f(Z + j\omega L_f) - j\omega M I_c$$

where Z is understood to mean $Z(jw)$.

Solving for I_c and I_f

$$I_c = \frac{E_c(Z + j\omega L_f) + E_0 j\omega M}{ZR_c + Zj\omega L_c + R_c j\omega L_f}$$

$$I_f = \frac{E_0(R_c + j\omega L_c) + E_c j\omega M}{ZR_c + Zj\omega L_c + R_c j\omega L_f}$$

The next step in relating the output voltage to the input voltage is to substitute these values for I_c and I_f in the following equation

$$E_0 = K\Sigma NI$$

$$E_0 = KN_cI_c - KN_fI_f$$

When this is done, the following equation results

$$\frac{E_0}{E_c} = \frac{KN_cZ + KN_cj\omega L_f - KN_fj\omega M}{ZR_c + Zj\omega L_c + R_cj\omega L_f + KN_f(R_c + j\omega L_c) - KN_cj\omega M}$$

This can be simplified by making use of the following relationships

$$N_cL_f = aN_cN_f^2 = N_fM$$

$$N_fL_c = aN_fN_c^2 = N_cM$$

therefore

$$\frac{E_0}{E_c} = \frac{KN_cZ}{ZR_c + Zj\omega L_c + R_cj\omega L_f + KN_fR_c}$$

This can be written as

$$\frac{E_0}{E_c} = \frac{K(N_c/R_c)}{1 + j\omega(L_c/R_c) + j\omega(L_f/Z) + K(N_f/Z)}$$

By making use of the relationships

$$L_f = aN_f^2,\ L_c = aN_c^2$$

the equation becomes

$$\frac{E_0}{E_c} = \frac{K(N_c/R_c)}{1 + a(N_c^2/R_c + N_f^2/Z)(j\omega) + KN_f/Z}$$

Since the feedback signal is obtained from a much higher power source than the control signal, the following relationship will generally hold over the frequency range of interest

$$\frac{N_f^2}{Z} << \frac{N_c^2}{R_c}$$

therefore

$$\frac{E_0}{E_c} = \frac{K(N_c/R_c)}{1 + j\omega(L_c/R_c) + K(N_f/Z)}$$

This can be written as

$$\frac{E_0}{E_c} = \left[\frac{(N_c/R_c)}{1 + j\omega(L_c/R_c)}\right] \times \left[\frac{K}{1 + K\left(\frac{1}{(1 + j\omega(L_c/R_c))}\right)(N_f/Z)}\right]$$

References

1. Principles of Servomechanisms (book), **G. S. Brown, D. P. Campbell.** John Wiley and Sons, Inc., New York, N. Y., 1948.

2. Servomechanisms and Regulating System Design (book), **H. Chestnut, R. W. Mayer.** John Wiley and Sons, Inc., New York, N. Y., 1951.

3. Response Time of Magnetic Amplifiers, **E. L. Harder, W. F. Horton.** *AIEE Transactions*, vol. 69, pt. II, 1950, pp. 1130–41.

4. Compensation of a Magnetic Amplifier Servo System, **H. H. Woodson, C. V. Thrower.** *Proceedings*, National Electronics Conference, Chicago, Ill., vol. 8, 1952, pp. 158–65.

◆

Discussion

M. Riaz (Massachusetts Institute of Technology, Cambridge, Mass.): The restrictions on the loading effects of feedback windings in a single-stage magnetic amplifier as imposed by the author to his generalized transfer function do not appear to be warranted. If one accepts the given linear representation of the magnetic amplifier as a single time-constant system, a formulation of the dynamics of this device may be obtained which, at no extra cost in complexity of the analytical expressions, includes the action of loading. The resulting transfer function may then be considered as truly generalized, since it is not penalized by assumptions which, in some practical cases, cannot always be justified. This generalized transfer function will also reveal the approximating conditions, independent of frequency, which are necessary to validate equation (3), and which are more precise than the condition $N_f^2/Z \ll N_c^2/R_c$. This transfer function is actually given by the author in the Appendix; it may be written in operational form as

$$\frac{E_0}{E_c} = \frac{K(N_c/R_c)}{1 + p[L_c/R_c + L_f/Z(p)] + K[N_f/Z(p)]} \qquad (6)$$

It should perhaps be clearly stated that the derivation of this relation is based on the assumption (quite justifiable in practice), of perfect coupling between control windings. The block diagram shown in Fig. 10 is a representation of the complete transfer function which, in contrast to Fig. 2, shows the mutual loading effects on both control and feedback windings. To conform to this exact formulation, corrective terms must be added to Table I, which concern only the equivalent time constants T_2 and T_3. These time constants corresponding to the first four feedback networks of Table I with negative feedback are shown in Table II.

I have used this generalized approach in connection with the study of feedback in d-c machines, which often exhibit transfer functions very similar to those of magnetic amplifiers. The method has been extended to cover the case of an n-winding machine, each winding being excited by a signal voltage through a linear passive network. The transfer relation is expressible in the form

$$E_0 = \frac{\sum_{r=1}^{n} KN_r \frac{a_r(p)}{Z_r(p)} E_r}{1 + p\sum_{r=1}^{n} \frac{L_r}{Z_r(p)}} \qquad (7)$$

where

E_r = signal voltage to winding r

$Z_r(p)$ = equivalent Thevenin internal impedance of field coupling network and signal source to which is added the resistance of winding r

$a_r(p)$ = open-circuit transfer ratio of coupling network connected in field r

Equation 6 appears then as a special case of the general form given in equation 7 when $n = 2$, $E_1 = E_c$, $Z_1 = R_c$, $E_2 = E_0$, $Z_2 = Z$, and $a_2 = 1$.

Block diagrams similar to Fig. 10 may directly be derived from the general transfer relations and, furthermore, with these representations, it is possible to include such nonlinear effects as those produced by saturation.

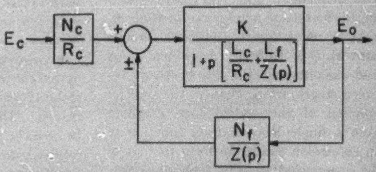


Fig. 10. Block diagram of complete transfer function, showing mutual loading effects on both control and feedback windings

Table II. Time Constants Corresponding to First Four Feedback Networks of Table I

Feedback Network	T_2	T_3
(A)	$\frac{T_A+(L_f/R_f)}{1+K(N_f/R_f)}$	
(B)	$\frac{T_1+T_A+(L_f/R_f)}{1+K(N_f/R_f)}$	$\frac{T_A}{1+K(N_f/R_f)}$
(C)	$\frac{T_1+T_A+[L_f/(R_1+R_2)]}{1+K[N_f/(R_1+R_2)]}$	$\frac{T_A+(L_f/R_2)}{1+K[N_f/(R_1+R_2)]}$
(D)	$T_1+T_A+KN_fC$	$T_A+(L_f/R)$

R. O. Decker: The restrictions on the loading effects of the feedback windings have been imposed as a matter of convenience in order to simplify both the block diagram and the expressions for T_2 and T_3. The restricting conditions are fulfilled in nearly all practical applications of magnetic amplifiers.

I am grateful to Mr. Riaz for pointing out the similarity between the transfer function of the magnetic amplifier and that of a d-c machine. However, the block diagram which Mr. Riaz proposes has more physical significance if it is drawn as shown in Fig. 11. Either block diagram gives the same over-all transfer function, but this one shows more clearly the dynamic relationship between the control voltage and the control current. For example, in a d-c generator this block diagram would indicate that the field current is not established instantaneously following a step change in control voltage.

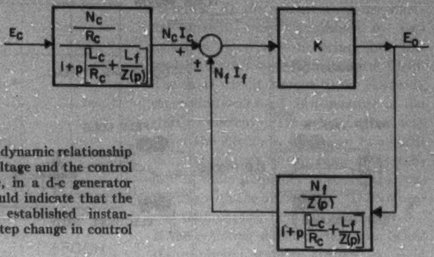


Fig. 11. Transfer function; dynamic relationship between control voltage and control current

---

Characteristics of the High-Current Argon Arc with Various Electrode Materials

J. W. DZIMIANSKI
ASSOCIATE MEMBER AIEE

T. B. JONES
MEMBER AIEE

Synopsis: A study was made of the fundamental properties of the high-current d-c arc in argon at atmospheric pressure, at currents of 25 to 100 amperes and for various electrode materials. Spectroscopic observations were made on the arc, and data for the voltage-current and voltage-arc length characteristics were taken for all materials tested. Anode melting rates were determined for copper, iron, and titanium. In addition, some information on minimum starting currents was obtained. These arc studies made it possible to develop a new empirical equation which depicts the electrical characteristics of the arcs investigated over a wide range of currents, including the positive slope region of the V-I curves. With this equation as a basis, reasonable energy and voltage balance relationships were developed and the relative values of the equation constants for the different materials were used for a comparison of their effects on the arc characteristics. In general, the arc is affected more by changes in gas composition than by changes in electrode materials. The melting rate data seemed to support a conclusion that in arcs of this type, over the current range studied, the anode drop is zero or extremely small.

THE Department of Electrical Engineering of The Johns Hopkins University has undertaken a series of investigations, sponsored by the Office of Naval Research, which are designed to lead to a better understanding of the high-current arc. Previous papers[1,2] have emphasized the effect of the gas by studying arcs with one type of electrode (tungsten) in various gases, namely, argon, helium, and numerous argon-helium mixtures. This paper reports another in this series of investigations, and its goal is to evaluate the effects of various electrode materials on arcs in one gas, namely, argon. Tungsten, tantalum, molybdenum, titanium, iron, and copper were chosen as electrode materials. They present as wide a variation of electrode properties as possible within the physical limitations of the experiments. Argon, rather than air, was chosen as the atmosphere to prevent chemical reactions and to limit the effect to that of a single gas. Furthermore, considerable interest in argon has been created by the expansion of the field of inert arc welding during recent years.

Description of Apparatus

A cross-section view of the arc chamber is shown in Fig. 1. It was constructed of brass with 0 ring seals at all demountable joints. Cooling was provided by a rectangular water jacket around the central portion, and two windows were built in to allow spectrographic and visual observations of the burning arc.

An oil diffusion pump and a mechanical fore pump were used to evacuate the unit to a pressure of 1×10^{-5} millimeters of mercury. Pressure was measured by thermocouple and Bourdon tube gages. Argon, 99.9 per cent pure, was supplied in gas cylinders at 2,000 pounds per square-inch pressure. A standard gas regulator valve reduced the pressure to 5 pounds' gauge, after which the argon was purified by passing through a refractory tube loaded with titanium sponge maintained at 900 degrees centigrade in a tube furnace. At this temperature, the titanium reacted readily with nitrogen, the principle impurity present, so that the argon admitted to the cell was of very high purity. A commercial high-frequency arc starter was used to initiate the discharge, and the arc current was supplied from a 300-ampere d-c arc-welding generator powered from a 220-volt 3-phase line. The spectrographic observations were made with a Baird 3-meter spectrograph. Arc voltage and current were continuously recorded on Esterline Angus recording meters.

Experimental Procedure

To achieve consistent results, the experimental procedures described in this section were followed rigorously. The desired gas purity was maintained in the cell by pumping it down to a pressure of 1 micron, flushing it with argon, re-evacuating it to 1-micron pressure, and then filling it to the working pressure, which was slightly above atmospheric. The cathode electrode was pointed to aid the cathode spot to anchor, then both electrodes were cleaned by polishing with a fine-grit emery paper, followed by a wash in hexane and then water. Observation of the arc voltmeter and projected image of the arc indicated that the arc voltage varied as the anode and cathode flames shifted at random intervals. To alleviate this condition, a set of coils was installed on the chamber. A magnetic field of approximately 500 gausses produced by these

---

Paper 54-510, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted October 16, 1953; made available for printing August 30, 1954.

J. W. DZIMIANSKI is with the Allis-Chalmers Manufacturing Company, Milwaukee, Wis., and T. B. JONES is with The Johns Hopkins University, Baltimore, Md.

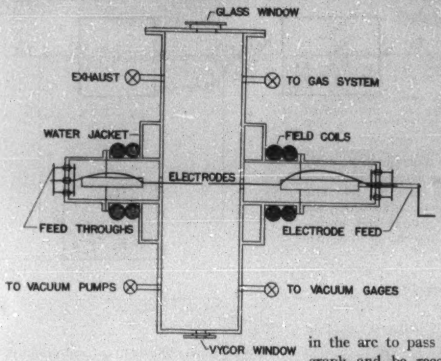


Fig. 1. Horizontal cross section of arc chamber showing all the major components

coils effectively pulled in the cathode and anode flames, lining them up with the electrode axes. In addition, the magnetic field constricted the plasma a slight amount, resulting in a burning voltage approximately a volt higher. A number of test runs and spot checks from time to time showed that the magnetic field in conjunction with shaping of electrodes could reduce the measured voltage variation for a given set of conditions to 0.2 volt.

Spectroscopic Measurements

The purpose of the spectroscopic observations was to determine how the electrode materials and the gas were distributed between the electrodes and also what types of ions (gas or metallic) were active in the important regions of the arc. The optical system was arranged so that a lens located between the arc and the entrance slit of the spectrograph projected an image of the arc on the slit and thus permitted light from selected regions in the arc to pass through the spectrograph and be recorded on the photographic plate. The data thus obtained could then be correlated with the electrical characteristics and the anode melting rates, to help clarify the internal process of the arc.

For these tests the electrode materials chosen were tungsten and molybdenum with an electrode separation of 1 inch and a current of 50 amperes. Argon was the gas in all cases. Spectrograms were taken beginning in a narrow region just in front of the cathode, followed by others across the arc plasma, with the final picture just in front of the anode. The first set of pictures was taken for a tungsten cathode and a molybdenum anode. Examination of these disclosed that the argon lines were extremely intense in front of the cathode, and much weaker, but present, throughout the rest of the space between the electrodes. Molybdenum lines appeared to be of about equal intensity throughout the entire discharge region, while tungsten lines appeared only in the vicinity of the cathode. Additional spectroscopic studies were made using another set of electrodes, with the polarity reversed to make the molybdenum electrode the cathode and the tungsten electrode the anode. This time the tungsten lines were found to be distributed throughout the entire region while molybdenum lines appeared only in the vicinity of the cathode. As before, the argon lines were very strong in front of the cathode and of lower intensity, but present, throughout the remainder of the arc. In general, these observations showed that for a 50-ampere arc up to 1 inch in length the anode material lines are found throughout the entire discharge, the cathode material lines are found only in the vicinity of the cathode, and the gas lines are present throughout the arc region with a large increase in intensity in front of the cathode. It seems reasonable to assume that in all cases where spectral lines of an element are found, ionization of that element has occurred and the degree of ionization is indicated by the intensity of the spectrum. It follows, then, that in the vicinity of the cathode all three of the elements (cathode and anode materials, and gas) supply charge carriers, while in the arc volume and in the region near the anode the gas and the anode material supply the positive ions.

Anode Melting

Estimates of the energy input to the anode were made by measuring melting rates of iron, copper, and titanium anodes. This was accomplished by weighing the anode before installation, burning the arc for a measured period of time at a fixed current, and then reweighing the electrode at the end of the experiment. The three melting rate curves determined from this information are shown in Fig. 2. At low currents the melting rates drop to zero very rapidly; the heat dissipated at low currents, from the electrode end by con-

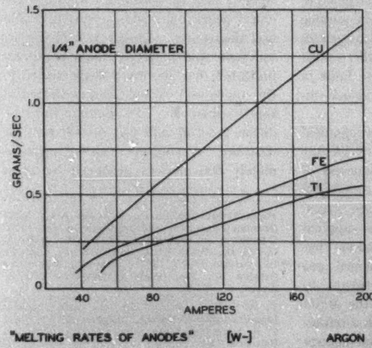
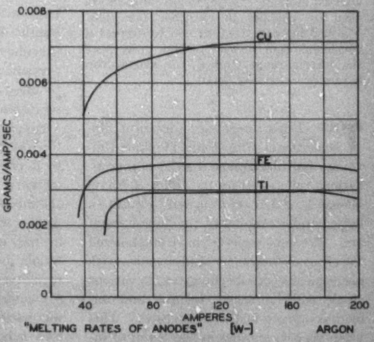


Fig. 2 (left). Melting rates of copper, iron, and titanium anodes in grams per second

Fig. 3 (right). Melting rates of copper, iron, and titanium anodes in grams per coulomb



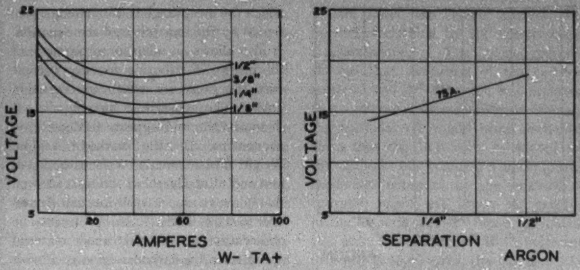


Fig. 4. V–I curves and V–S curve of arc between tungsten cathode and tantalum anode. Electrode diameter is 1/4 inch and arc lengths are as marked beside each V–I curve

duction, radiation, and convection, is a measure of what must be supplied before melting starts. In the intermediate current range, from a little above the start of melting up to 200 amperes, the curves seem to be linear. At higher currents they start to level off, probably because there is such a high energy input that insufficient heat is conducted through the electrode, resulting in the superheating of the melted material with a consequently greater radiation and convection heat loss. An interesting modification of Fig. 2 was secured by dividing the melting rate by the current, giving a new melting rate in grams per coulomb. Fig. 3 effectively shows that, once melting starts, the rate per unit of charge passed through the anode quickly reaches an almost constant value over a wide range.

The sources of heat for the anode may be enumerated as follows:

1. Heat of condensation of electrons, $I\phi$.
2. Thermal energy the electrons possess, $2kT$.
3. Radiation from the cathode and plasma.
4. I^2R loss in the anode.
5. Kinetic energy imparted to electrons by an anode drop.

Calculations show that the largest percentage of the anode heat, at least 80 per cent, comes from the heat of condensation of electrons and approximately 15 per cent may be derived from the thermal energy which the electrons possess, assuming a plasma temperature of 6,000 degrees Kelvin. Radiation from the cathode and plasma and I^2R loss add only a very small amount in this case. Since processes 1, 2, 3, and 4 always take place, and since they account for over 95 per cent of the heat input to the anode, one must conclude that no appreciable heat is derived from an anode drop. This indicates that the anode drop is zero or very small.

Copper may be used as an example

Heat input per coulomb from melting rate = 1.03 calorie
Heat of condensation per coulomb = 1.02 calorie
Thermal energy per coulomb of electron = 0.20 calorie

The heat input to the anode necessary for the slight evaporation of copper that occurred was much less than the experimental error of measuring the anode melting rate and therefore has not been included in the tabulation. With due allowance given for the lack of exact knowledge of the work function for the anode surfaces and plasma temperature, the heat developed at the anode, to at least the first approximation, can be considered as attributable to the heat of condensation of electrons, $I\phi$.

Arc-Starting Current

Because of difficulties encountered at times in initiating an arc between test electrodes, a number of checks were made on starting currents between several electrode materials in argon, helium, and air. Doan[3,4,5] seems to have been the first to report the inability to start an arc in very pure inert gases at low currents. He succeeded in initiating such arcs only at high currents. In the present work, arc-starting currents were determined by two methods. One was to bring the electrodes together and then separate them. The short-circuit current before electrode separation was used as a measure of the starting current. The other method was to set the electrode separation to a small value (1/13 inch in this case) and establish a radio-frequency discharge across the gap with the arc starter. D-c voltage was then applied between the electrodes. If the resulting direct current was below the minimum starting current, no rigid cathode flame appeared, and the discharge was immediately extinguished when the radio-frequency voltage from the arc starter was removed. The values of starting current determined by the radio frequency arc starter method were about 10 per cent lower than those found by the short-circuit method. All the measurements were taken with a generator open-circuit voltage of 60 to 70 volts and with the stabilizing magnetic field turned off. No self-sustaining arc would form below the minimum starting current but, once formed, the current could be decreased to as low as one-third of the starting value before the arc was extinguished. For 1/4-inch tungsten electrodes the starting current was about 20 amperes in argon, about 100 amperes in helium, and about 2 amperes in air. Reduction of the cathode diameter reduced the starting current. A 1/16-diameter tungsten cathode required only 8 amperes in argon. Variation of the anode diameter did not seem to affect the starting current.

Voltage-Current and Voltage-Separation Measurements

Probably the arc characteristics most important to the electrical engineer are the variation of voltage across the arc as

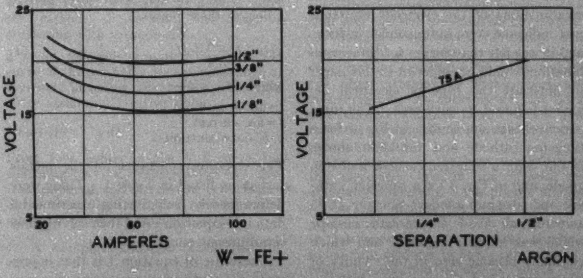


Fig. 5. V–I and V–S curves of arc between tungsten cathode and iron anode. Electrode diameter is 1/4 inch and arc lengths are as marked beside each V–I curve

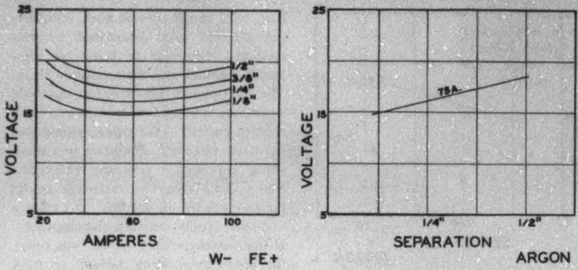


Fig. 6. V–I and V–S curves of arc between tungsten cathode and titanium anode. Electrode diameter is 1/4 inch and arc lengths are as marked beside each V–I curve

the current is changed, while maintaining constant electrode separation (V−I curves); and the variation of voltage across the arc as the electrode separation is changed, while maintaining constant current (V−S curves). Considerable effort was expended to secure curves as precise as possible by following all of the steps outlined in the experimental procedure. With a tungsten cathode V−I and V−S curves were obtained for anode materials of copper, iron, titanium, molybdenum, tantalum, and tungsten. Then, with tungsten as the anode, curves were obtained for tantalum, molybdenum, and tungsten as cathodes. Also, for comparative purposes, one set of curves was determined for tungsten electrodes in helium. The V−I curves were obtained at electrode separations of 1/8, 1/4, 3/8 and 1/2 inch for currents of 25 to 100 amperes.

Accurate curves could not be obtained for copper, iron, or titanium cathode materials because of the rapid motion of the cathode spot over the electrode surface. The voltage and current values were continuously fluctuating, and it was difficult to keep the arc burning for more than a few seconds at a time, even with the magnetic field.

The curves obtained for the various combinations of the electrode materials just indicated were quite similar in form, and it was felt that curves for three combinations would be sufficient in this paper to illustrate the type of electrical behavior followed by such arcs. The three curves chosen are shown, in Fig. 4 for a tungsten cathode and tantalum anode, in Fig. 5 for a tungsten cathode and iron anode, and in Fig. 6 for a tungsten cathode and titanium anode. Starting at 25 amperes, all V−I curves had a slope which was initially negative, and which gradually became zero in the vicinity of 60 amperes, and then positive with increasing current. The V−S curves were in all cases essentially straight-line functions. The slopes of these lines were the plasma gradients, and they varied from 9.5 volts per inch for an iron anode to 12.5 volts per inch for a titanium anode. The effect of the various electrode materials could not readily be determined from visual studies of the voltage-current and voltage-separation curves because the curves for the different materials were too nearly alike. An equation representing the electrical characteristics of the arc as depicted by the curves was first developed and then analysis of the various parts of this equation brought out the electrode influence to the extent that it could be determined.

Since the earliest studies of the arc, investigators have developed empirical equations describing the relationship between voltage, current, and the length of the arc discharge. Most of these equations covered only the low-current range below 20 amperes. A search of the literature disclosed no equation that would agree with the experimental results obtained here. Upon close examination of data equation 1 was found to satisfy the experimental curves over their entire range of 25 to 100 amperes. Furthermore, some measurements taken at 200 and 450 amperes indicated that equation 1 held at these points.

$$e = a + bl + \frac{c}{I} + dI \qquad (1)$$

where

l = arc length
I = arc current
a, b, c, d = constants

The equation constants for the metals tested are listed in Table I. These were determined by substituting experimental data into equation 1 and then solving four simultaneous equations.

The value of equation 1 is that it presents a quantitative comparison for any given arc of the four basic elements of which the arc voltage is comprised, represented by the four terms of the equation. It also allows an analysis to be made of the energy distribution for any current over the range for which the equation holds. The constant a has the dimensions of volts and appears to depend on phenomena in the cathode region. Weight measurements on the cathode before and after several of the tests showed that some anode material was transferred over to the cathode. Also spectrographic observations showed that anode material appeared in the cathode region even with fairly large arc lengths. The constant a is probably determined by a composite-excitation potential of the mixture of gas and metal vapor existing in the cathode region, and also by the work function of the cathode surface contaminated with anode material and gas.

The constant b represents the gradient in the column in volts per inch, and measurements showed it to be constant for any electrode separation of from nearly zero to at least 1.25 inches. Also, it did not vary as the current was varied over the range of 25 to 100 amperes. The constant b is an indication of energy losses in the arc plasma.

The constant c accounts for the negative slope portion of the voltage-current characteristics, and had the dimensions of watts. This constant seems to be associated with the dissipation of heat in the cathode electrode. The brightness of the cathode remained nearly constant when the current was varied, indicating constant cathode power input, while the positive electrode showed considerable change of brightness with change in current. Also from Table I it can be seen that for a tungsten cathode the value of c is consistently in the neighborhood of 225 watts, regardless of the material of the anode. On the other hand, when the cathode material is changed to molyb-

Table I. Calculated Equation Constants

	Volts a	Volts per Inch b	Watts c	Ohms d
Tungsten Cathode in Argon				
Anode				
Tantalum	3.8	12.0	255	0.074
Molybdenum	4.6	12.0	260	0.052
Iron	5.7	9.5	210	0.070
Tungsten	6.5	10.3	235	0.047
Titanium	7.6	12.5	186	0.049
Copper	9.5	11.5	194	0.029
Tungsten Anode in Argon				
Cathode				
Tungsten	6.5	10.3	235	0.047
Molybdenum	7.4	11.2	107	0.036
Tantalum	5.6	12.5	186	0.031
Tungsten Electrodes in Helium	14.8	44.8	445	*

* The arc in helium was not measured at sufficiently high currents to determine this constant.

denum, the value of c falls to 107 watts. This indicates that it is primarily a characteristic of the cathode.

The constant d has the dimensions of resistance and determines the value of the positive slope of the voltage-current curve. At first this constant was thought to represent spreading resistance from the nonuniform flow of current into the end of the anode or cathode. However, energy dissipated by this term as I^2R loss in either cathode or anode did not show up. At the higher currents this energy would have been sufficient to melt even the high-temperature cathodes like tungsten (which did not melt) and also to produce an anode-melting rate far above that which was actually measured. It is probable that d is part of the cathode drop and represents energy dissipated in the cathode flame, near the cathode surface but not in the body of the metal.

An estimate of power dissipated in the various portions of the arc was made by multiplying the arc equation 1, through by the current I so that each term represents power in watts.

$$eI = aI + bII + c + dI^2$$

The bII term can be accounted for readily as it is the power consumed in the plasma, and from the melting rates of the anodes we know that approximately $I\phi$ is dissipated there. The term c appears to be power dissipation in the cathode electrode. The remainder $(aI+dI^2)-I\phi$ is considered to be dissipated in the cathode region. Using information in Table I, the power distribution for a 50-ampere arc between tungsten electrodes separated 1/2 inch may be computed as follows:

anode
$I\phi = 50\times 4.8$ $= 240$ watts

plasma
$bII = 10.3\times 1/2\times 50$ $= 257$ watts

cathode electrode
c $= 235$ watts

cathode region
$(Ia+dI^2)-I\phi = 50\times 6.5 +$
$0.047\times(50)^2 - 240 = 203$ watts
Total 935 watts

The voltage distribution for the 50-ampere arc appears to be as follows:

cathode drop

$$a+\frac{c}{I}+dI = 6.5+\frac{235}{50}+0.047\times 50 =$$
13.6 volts

plasma drop
$bl = 10.3\times 1/2$ $=$ 5.1 volts
Total 18.7 volts

This power and voltage analysis was somewhat clarified by the spectroscopic measurements. The spectrograms showed that excitation of the cathode material and of the argon gas occurred to an intense degree directly in front of the cathode. It would seem probably, then, that the 13.6 volts calculated for the cathode fall represents a composite excitation potential of the mixture of tungsten vapor and argon gas existing in the cathode region. Since the ionization potential of tungsten is 8.1 volts and of argon is 15.69 volts, the value of 13.6 volts appears reasonable.

Excitation lines of the anode material and of argon are found throughout the arc region, indicating that ionization of these two components provides ionic charges throughout the plasma. Actually the metallic ions of the anode or cathode materials are not necessary to sustain the arc, the surrounding gas being sufficient. This is demonstrated in extremely long arcs of 1 foot or more in length. The metallic ions are merely byproducts of the heat associated with the fundamental electron emission and condensation processes at the cathode and anode respectively. In short arcs of the type studied here, the metallic ions are of considerable importance, and their low ionization potential causes a substantial reduction in the arc voltage for a given current.

Conclusions

1. Electrical characteristics of the arc between metal electrodes in argon gas can be reproducibly determined to within a fraction of a volt when appreciable vaporization and melting of the cathode does not exist. This reproducibility is achieved when the negative electrode is pointed to anchor the cathode spot and a coaxial magnetic field is used to control the arc column.

2. There is a minimum current below which an arc will not start in pure argon or helium between clean electrodes. The starting currents appear to depend upon the cathode dimensions, increasing or decreasing with cathode size.

3. The following equation for the V–I curve of the arc holds over a wide range of currents, including the region where the curve assumes a positive slope

$$e = a + bl + c/I + dI$$

From four points in the V–I curve, the constants may be evaluated for any electrode material, and thereby the performance of the arc can be determined over the entire range for which the equation holds.

4. A complex situation exists at the cathode.

5. The plasma gradient for electrode separations up to 1/2 inch is essentially constant; measurements up to 1¼ inches indicate constancy of the gradient up to this separation also.

6. Anode-melting rates indicate that the condensation of electrons supplies most of the energy to the positive electrode. An anode drop does not seem to exist in this type of arc.

7. Spectroscopic studies show that the anode material and enclosing gas are ionized throughout the arc region and are the source of positive ions everywhere in the arc except directly in front of the cathode. The major voltage drop in the arc occurs directly in front of the cathode, providing the power necessary for electron emission, for intense ionization of the enclosing gas, and for some vaporization and ionization of the cathode material.

8. The arc voltage did not change as much with electrode materials used as with different gases used, such as argon and helium. The ionization potential of all metals tested differed, however, by only 1.3 volts, whereas the ionization potential of helium is 8.8 volts higher than that of argon. Other factors, such as ionic mobilities, enter the picture, and additional studies along these lines should further clarify the effect of the enclosing gas and the electrode materials on the behavior of the high-current arc.

References

1. THE ELECTRIC ARC IN ARGON AND HELIUM, **T. B. Jones, Merrill Skolnik, W. B. Kouwenhoven.** *AIEE Transactions*, vol. 72, pt. II, Mar. 1953, pp. 16–21.

2. CHARACTERISTICS OF THE HIGH-CURRENT TUNGSTEN ARC IN ARGON, HELIUM AND THEIR MIXTURES, **Merrill Skolnik, T. B. Jones.** *Journal of Applied Physics*, New York, N. Y., vol. 23, 1952, pp. 643–52.

3. ARC DISCHARGE NOT OBTAINED IN PURE ARGON GAS, **G. E. Doan, J. L. Meyer.** *Physical Review*, New York, N. Y., vol. 40, 1932, pp. 36–39.

4. ARCS IN INERT GASES, **G. E. Doan, A. M. Throne.** *Ibid.*, vol. 46, 1934, pp. 49–52.

5. ARCS IN INERT GASES, **G. E. Doan, W. C. Schultze.** *Ibid.*, vol. 47, 1935, pp. 783–34.

◆

Discussion

Robert C. McMaster (Batelle Memorial Institute, Columbus, Ohio): The authors are to be congratulated upon their unusually clear and effective presentation of excellent experimental techniques and results. The experimental conditions provided a degree of control and consistency of measurements not often attained in high-current arc measurements. In consequence, the data and conclusions appear more clear-cut and valid than are usually obtainable from such experiments.

The analyses of the heat economies at the electrodes and in the plasma are unusually interesting. The plasma studies appear to confirm the observations of prior investigators. It would be of interest to determine how closely these observational data conform to Suit's law of similitude (conservation of energy across the plasma boundary, by equating heat produced in the plasma to heat transferred radially outward across the boundary). The facts that heat delivered to the anode appears to be supplied, for the most part (at least 80 per cent) from the heat of condensation of electrons, and that no appreciable heat is derived from the anode drop, deserve particular consideration. It may prove to be desirable to reexamine present concepts of the design of welding arcs and metal-melting furnace arcs in terms of this finding.

The volt-ampere relationship (equation 1) and the corresponding justification of its terms also offer considerable food for thought, particularly with respect to the constant *d*. If energies of the order of 100 watts (10 per cent of the arc energy) are dissipated here, some study of the cathode flame phenomena and the control of this heat loss (for example, by directing it back onto the work in arc welding) may be justified. It is not immediately evident, however, how the cathode flame, which may be in parallel with the arc column (viewed as an electric circuit) contributes directly to the arc voltage drop. In terms of energy losses, however, this approach deserves consideration.

J. W. Dzimianski and **T. B. Jones**: Mr. McMaster's discussion is appreciated. The various physical processes taking place at the cathode were not resolved in our investigation; however, the existence of a complex situation in the vicinity of the cathode was clearly indicated. The use of the coaxial magnetic field in our experiments fixed the geometry of the arc so that the cathode flame was coincident with the arc column. In the case where the arc column is not coincident with the cathode flame, the portion of the arc drop represented by the constant *d* part of the volt-ampere relationship (equation 1) should be attributed to the arc region in the vicinity of the cathode rather than to the cathode flame.

Some Recent Advances in the Economy of Routing Calls in Nation-Wide Toll Dialing

IMRE MOLNAR
MEMBER AIEE

Synopsis: In the first part of the paper some of the important considerations in nation-wide toll switching are discussed. In the second part an arrangement in a new intertoll switching system is described which permits considerable simplifications in the equipment for the routing of toll calls.

Nation-Wide Toll Switching

THE unprecendented growth in telephone stations and the demand for toll telephone service in the United States and Canada since 1945 has created a strong economic pressure towards mechanization of toll operation on a continent-wide scale. The ultimate goal of complete mechanization can naturally be achieved only in progressive steps spread over several decades, the first phase aiming at operator toll dialing, and the second phase at subscriber toll dialing, with both phases overlapping at times.

Toll dialing has been practiced on a limited regional scale in the United States for many years, by both operators and subscribers, and among both independent and bell system telephone companies. However, it is only recently that the telephone industry has devised and embarked upon an integrated plan for toll dialing which will embrace all of the United States and Canada. At the present time, about half of the long-distance toll calls are dialed by the originating operators, while the proportion of subscriber-dialed long-distance calls is still negligibly small. In short-haul toll traffic the great majority of calls are dialed either by operators or by the subscribers.

Integrated subscriber toll dialing networks have been serving for many years in several European countries. Such networks have been working in Southern Germany for over 25 years; others are in operation in Belgium and Holland, while toll traffic is now dialed by subscribers over all of Switzerland. Much can be learned from their collective experience, but however impressive their technological progress may be, an integrated toll dialing plan for the whole North American continent has required many new approaches, due partly to the types of service given here by telephone companies, partly to the telephone habits of the public, and partly to the very magnitude of the problem. It should be borne in mind, for instance, that the area of Switzerland is two-thirds that of West Virginia, that the total number of telephones there is about the same as in the city of Detroit, and that the number of telephones in the United States and Canada is almost twice that of the rest of the world together.

To establish a toll dialing plan involves a multitude of prerequisites, of which the following are the more obvious:

1. Toll dialing circuits. Such circuits must naturally be capable of transmitting dial pulses and other supervisory signals. They must be reliable, reasonably free from signal distortion, and immune from disturbances which might cause false signals or service disruptions, particularly when used in a subscriber toll dialing network.

2. Automatic central office equipment. Toll dialing assumes that the central office at the point of destination is converted to automatic operation, though occasionally such calls may be terminated in call indicator types of manual offices. For subscriber toll dialing the originating office must likewise be automatic and the subscribers' dials must be of a suitable type.

3. Automatic intertoll switching equipment. At intermediate and terminating points, automatic switching equipment must be provided for through switching of toll calls and for switching toll calls into local automatic equipment as well as to special service equipment.

4. Transmission means. The toll circuit plant must meet the requirements of the prescribed transmission standards regardless of the various types of intertoll circuits, intertoll switching equipment, and the various routes and number of switching points for the same destination.

5. Numbering plans. Suitable numbering plans must be established for all subscribers who are to receive toll calls over the intertoll dialing network, and suitable numbering plans must be established for routing and switching such calls through the toll plant.

6. Circuit provisions. With automatic routing of toll calls a higher speed of service and therefore a more liberally engineered toll plant is required than with purely manual switching. When toll calls are to be dialed by subscribers, the provision of the toll plant must approach the virtual no-delay performance of local switching.

7. Toll ticketing means. For subscriber toll dialing means must be provided for

---

Paper 54-503, recommended by the AIEE Communication Switching Systems Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 21, 1954; made available for printing August 23, 1954.

IMRE MOLNAR is with the Automatic Electric Company, Chicago, Ill.

automatic registration of the toll charges and means for transferring these charges to the subscriber's account. The introduction of such automatic toll ticketing equipment makes it economically feasible to introduce equipment to provide special routing and switching functions, over and above those required for operator toll dialing.

In addition to these items there are a large number of others, less obvious but nonetheless important. They may be of a technical nature such as the provision for intercept service, or purely administrative such as public relations, or a combination of both such as directory information. Item 2 does not require further discussion; item 1 has been extensively treated in technical literature and there will, no doubt, be further improvements before completely satisfactory techniques are firmly established. Item 4 is in a state of fluidity and great advances both in general standards as well as in improved means can be expected.[1] Items 3 and 5 are the subject of this paper which will deal with new and simplified approaches; item 6 will be briefly touched upon. Item 7 is outside the scope of this paper inasmuch as it concerns only the ultimate objectives of subscriber toll dialing and these can be superimposed upon the structures discussed here.

The numbering plan, as devised by the Bell system, is based on the assumption that ultimately every telephone subscriber in the United States and Canada will be assigned a number which will be different from the telephone number of any other subscriber. Such numbers will uniformly consist of ten digits, the first three of which will designate the numbering plan area, the next three the central office, and the last four the subscriber's number in the central office. It is further planned that ultimately for routing of calls outside the calling subscriber's own numbering plan area all ten digits will be required; that for calls within that numbering plan area only the last seven digits will be dialed, though in some cases seven digits will suffice for calls between certain offices in two adjacent areas having a high community of interest; and that calls within smaller communities may also be established by the dialing of the last five or four digits only.

The United States and Canada, taken together, are divided into about 90 numbering plan areas with other areas being added where the present subdivision proves to be inadequate. A state or province includes at least one whole numbering plan area while many states or provinces have to be subdivided into several—as many as six—independent numbering plan areas. Each area code has 0 or 1 as its second digit, while the first and third digits are numbers from 2 through 9; therefore, with this scheme, it is possible to increase the number of areas to 128. By this numbering arrangement, in contrast to the central office codes, the toll switching apparatus can distinguish between calls destined to points within or outside of the numbering plan area.

To each central office within an area a distinct 3-digit code is given consisting of the numerical equivalent of the first two letters of the central office name plus one numerical digit from 1 through 0. The same central office codes may be reused in many or all of the 90 numbering plan areas. At present about half of the total telephones are operated on this so-called 2–5 numbering basis. The maximum theoretical number of central office codes per area is thus 600, excluding the 40 codes set aside for radio-telephone service. The practically available number of codes, however, is considerably less, for a block of codes are required for the so-called toll center codes, further codes are used up when conflicts are to be avoided where 7-digit dialing is required between adjacent numbering plan areas, other codes may have to be omitted because of undesirable central office names, and codes may also be sacrificed for the sake of economies in trunking and switching equipment. At present it is planned that a separate code will be required for every central office even though it may serve only a few stations, particularly if the office represents a distinct toll rate center, since it is assumed that it will be practical only for a machine to distinguish toll rates on a 3-digit basis.

The last four digits represent the station number within a central office assuming central office units from a few stations up to 10,000. There are a large number of central office units in service having 5-digit station numbers which, unless blocks of central office codes are to be used up, result in 11-digit toll numbers. Such number assignments are common in larger terminal-per-line central offices, and also in high-multiple manual boards. The future of these 11-digit numbers is still uncertain. It is felt that when subscriber toll dialing is introduced on a wide scale it would be confusing for subscribers to dial 7-digit or 10-digit numbers in most cases (and perhaps 5-digit local numbers as well) and eight or 11 digits for some destinations. On the other hand, if squandering of the already limited number of central office codes is to be avoided, the elimination of such numbers might in many cases be unduly costly because this requires conversions in the local switching equipment. In common control types of intertoll switching systems and toll ticketing equipment, mixing of 10-digit and 11-digit numbers causes some, though not excessive, circuit complications which should preferably be avoided. In direct pulse controlled intertoll switching systems, the number of digits is inconsequential and in some cases such mixed numbering has advantages in flexibility and economy of trunking. How this problem will ultimately be disposed of cannot now be predicted; with operator toll dialing mixed numbering is no problem at all, and with the limited extent of subscriber toll dialing which will be probable in the next few years such 11-digit destinations can be excluded and left to operator completion if desired. However, in general there is a tendency by engineers to underestimate the competence and willingness of the general public, just as years ago it was thought that automatic telephone systems were impractical because subscribers would be unable to dial their calls and would always need operator assistance.

The theoretical capacity of the North American numbering system is 768,000,000 station numbers, which is 14 times the the present total stations in the United States and Canada. We have seen, however, that for practical purposes this picture is misleading and we must look at it from a different angle. The theoretical maximum number of office codes which can be accommodated by this plan is 76,800 but practical considerations will limit this to a much lesser number, say on the order of 50,000. Since the present number of central offices in the United States and Canada is about 22,000 there is still a comfortable 150-per-cent margin, but only if the patterns for both local and toll trunking are very carefully laid out in an integrated plan which avoids waste in central office codes and this will be possible in some cases only at some additional cost in plant and equipment. Actually, with the present 90 numbering plan areas the theoretical maximum number of office codes is 54,000 which considering practical limitations gives us a less comfortable margin to work with. This problem is of particular importance to independent telephone companies since two-thirds of the central offices are outside the Bell system.

We have assumed in the preceding discussion that a universal numbering plan is a prerequisite for toll dialing which means that the same number is dialed to reach a certain station no matter where on the continent the call originates;

however, some might question the validity of this assumption. Actually, even the ultimate plan does not contemplate such uniformity. A subscriber in a small town should and will be able to reach a station in his community by dialing a 4-digit or 5-digit number though he can also reach the same station by dialing the full seven digits; but if he calls from a nearby town he must dial the full number. Furthermore, a subscriber in Chicago, e.g., will reach any station in Chicago or in any of its suburbs, or any one in the same numbering plan area, or any one in Gary, Ind., another numbering plan area, by dialing a 7-digit directory number, but must prefix one of the several Illinois area codes if he calls a station in Illinois outside his own area. It is difficult to see how this can be avoided by any practical system. But it might as well be realized that subscribers and private branch exchange operators must have available a certain amount of routing information for making long-distance calls even though this will be very simple; and for frequently called destinations which represent a large portion of toll traffic the routing information can readily be memorized.

In some of the European countries where subscribers dial their toll calls, extensive use is made of routing codes consisting of various numbers of digits, apparently with satisfactory results. However, for subscriber toll dialing a universal numbering plan, in the sense described in the preceding paragraph, appears to be the more desirable solution if it can be accomplished at a reasonable cost. "Reasonable" is a loose word to use in the guidance of engineering design, and universal numbering does cost some money. Subscriber toll dialing is still in its infancy and what its ultimate cost will be remains to be seen. However, huge investments in toll plant are already being made towards that end, and in spite of generally rising price levels and increasing operating costs it has been possible to reduce long-distance toll rates. While this fact in itself is insufficient to justify every phase of the program, it provides at least some evidence that the cost for the over-all approach is not yet unreasonable.

In operator toll dialing either direct point-to-point circuits or directing codes are used, followed by the subscriber's directory number. The directing codes might be arbitrary to the point of being governed by the trunking pattern of the switching equipment; they might resemble the national codes though they were chosen by similar considerations; or they might coincide with the national codes of the universal numbering plan. During the next few years operators will be able to dial more and more calls by using the national codes, but it is not expected that this program will be substantially complete for at least another 10 years. Whether all of the arbitrary routings by operators can ever be eliminated or whether this would necessarily be desirable is open to debate. When nation-wide toll dialing is discussed, instinctively cross-country long-haul traffic is thought of and it is apt to be forgotten that the community of interest usually declines with distance and that the large volume of toll traffic remains close by, often within the same numbering plan area. There are at the present about 2,500 toll centers and only about 70 of these according to present plans will become control switching points. The operators in the other toll centers will handle large volumes of traffic over direct trunks and alternate routes not passing through control switching points. Advantage is taken at such locations of the inherent simplicity and flexibility of Strowger-type intertoll switching equipment without common control to build up multiswitched connections at low cost and at maximum speed. Operators quickly memorize the frequently used routing codes, while others are readily found in their flip-flop references or keyshelf bulletins to which they must often refer for other purposes also. With careful planning of code assignments these routing codes can in many cases coincide with the national numbering, in which cases further simplification in switchboard operation will result. Gradually a point of diminishing returns may be reached and the remaining items of traffic handled by arbitrary routing assignments rather than by added equipment with its increased complexity and cost.

In addition to routing calls to telephone subscribers, the intertoll switching plan must provide for routing calls to various operators and to various plant test facilities. At present it appears that such routings, if feasible, should be restricted to calls dialed by operators. Special number assignments are made for such service codes consisting of three or four digits or, in some cases, five digits, the special distinction for such codes being that all of them have 1 as the first digit. When such service calls are routed to toll centers, a so-called toll center code is prefixed to the service codes. The toll center codes are similar to the regular central office codes; in multioffice areas a toll center code is assigned which does not conflict with any office code but which preferably ends with 0; in smaller toll centers the code may be the same as the central office code. In toll centers which are not control switching points, the toll center codes are omitted and such service calls are switched upon the first digit 1 of the service codes.

With the introduction of toll dialing by operators, it became apparent that in order to obtain the maximum benefit from this method a higher speed trunk service is required than with the previous ringdown operation, partly because delays at intermediate switching points with consequent repeated attempts were to be avoided. For most long-distance routes the objective was a busy-hour trunk speed of 30 seconds, with some routes engineered at a better and a few others at a worse grade of service. Special provisions were devised at switching points to indicate to the originating operators the state of congestion and to expedite traffic, but in general more liberally engineered trunk groups were required to obtain the desired service objective. On the other hand, toll dialing permits the elimination of unprofitable direct circuits by the substitution of more switching points than are permissible with manual switching; thus, by combining several independent traffic paths into one circuit group, a higher efficiency per trunk is obtained and the cost of increased speed of service is reduced.

With subscriber toll dialing, a still better grade of service is required since if the customer is too often blocked on the route to his destination, he will get discouraged and pass the call to an operator. Many years of experience in local operation provides good indication on the grade of service customers demand, and the engineering of toll circuit provisions for subscriber toll dialing can benefit from this experience. To approach the interlocal grade of service in toll operation would be prohibitive with the present cost of toll circuits and it could well be that long-distance subscriber toll dialing would have been impossible if multiswitching in dial operation had not offered increased economies.

Toll operation with alternate routing and multiswitching has been practiced probably as long as congruous toll networks have existed. Automatic alternate routing in local operation has been practiced for many decades even when it was not known by this name. Operator toll dialing is possible with or without automatic alternate routing. The originating operator can see at a glance when the principal route is busy and can choose at once one of the prescribed alternatives, and can do this at no loss of speed and at a lower cost than by automatic means. Visual alternate routing has the advantage of flexibility by permitting instantaneous changes in the assignment of alternate routes when traffic or plant conditions require such changes on short order, without administrative problems or interference with the switching equipment. It has the further advantage that the operator knows the route the call has taken in case this information is needed for arriving at toll settlements. Alternate routing can similarly be treated by the originating operator at intermediate switching points though automatic alternate routing at such points has some definite advantages. Alternate routing of calls for subscriber toll dialing without automatic means appears to be impractical.

Considerations governing automatic alternate routing from the traffic engineers' point of view were described in a recent article.[2] It appears that under the existing conditions substantial economies can be achieved with multiswitching and automatic alternate routing for the grade of service required in subscriber toll dialing as compared with a system without alternate routing; however, whether for the volume of traffic in, say, 10 years from now, these advantages will remain the same or become greater or less seems at the moment to be unpredictable. Even the magnitude of savings achievable today cannot be reliably evaluated because of the inadequacy of statistical methods and lack of standards on which comparisons can be based; such evaluation must await the appearance of sufficient empirical data.

The uncertainty of the situation to be faced 10 years from now is even more pronounced if we consider that we have no idea about the degree of progress which will be achieved in technical development, particularly in regard to further cost reductions for transmission means. However, judging from what has happened in the past 10 or 15 years there is every reason to believe that such progress will be considerable. The whole concept of transmission media has changed. Not so long ago the cost of a toll circuit was largely governed by so many miles of copper and lead and so many poles and crossarms. Since then the importance of these items has become progressively reduced, because they have been replaced by concentrated components in the form of filters, condensers, vacuum tubes, etc. High usage circuits are established between two toll centers when such circuits are more economical than any other available route so it might well be that at some future date the cost of toll circuits will compare favorably with the switching cost and to provide direct circuits will become more attractive than alternate routing, or at least the proportion of traffic which is to overflow from the high usage circuits may be reduced.

These considerations do not necessarily lead to the conclusion that automatic alternate routing is not justified or that it will shortly become obsolete or that engineering developments for such systems are unjustified. There can be no doubt that in many cases automatic alternate routing represents a considerable saving and that this will remain true in the foreseeable future. Therefore there is a real need for simple automatic routing systems at the lowest possible cost. Besides being simple and inexpensive, they should be flexible and self-contained with the provision that they can be added or removed whenever conditions so require. Such an arrangement as part of a new intertoll switching system was recently described.[3]

Intertoll Switching System

The preceding discussion has dealt with some of the most important aspects of nation-wide toll dialing; the remainder of the paper will describe simplifications in the routing of toll traffic as incorporated in the recently developed intertoll switching system referred to in the previous paragraph. Since at the present time and for some years to come the great majority of toll calls will be extended by operators, the discussion will relate mainly to operator toll dialing. The routing of subscriber dialed toll calls is incorporated in the automatic toll ticketing equipment which can be superimposed on this arrangement.

For the automatic switching of toll traffic in this country either Strowger type or crossbar type of switching equipment is employed. Both systems are now fully developed to provide all features required for intertoll switching. The Strowger equipment might be straightforward dial pulse controlled, wholly of the common control type, or a combination of the two with common control being used only as far as it is economically justified. Though dial pulse controlled crossbar switching systems have been used abroad to a limited extent, the developments here and abroad during the past 10 to 15 years have demonstrated that the full potentialities of crossbar equipment can be exploited only by complete common control. Therefore all crossbar equipment used in the United States is so designed. Strowger equipment has been well known to telephone people for many years, while large-scale use of crossbar systems is comparatively new to most of them, even though sporadic applications of crossbar equipment have appeared for at least 35 years. Therefore recent technical publications have to a large extent been devoted to descriptions of the less familiar features of crossbar equipment, while the advances in Strowger techniques as applied to intertoll switching have apparently been taken for granted. The planning of a large intertoll switching network employing Strowger-type equipment exclusively, and taking full advantage of its potential economy, simplicity, and flexibility, will now be described. Because there are many areas with similar characteristics, the discussion will be best illustrated not by hypothetical examples but rather by factual data, projected forward

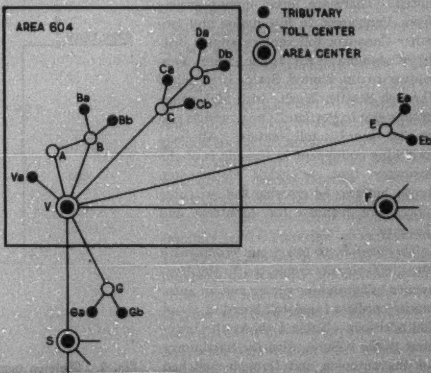


Fig. 1. Basic toll routing plan

Table I. Busy-Hour Traffic Volume

	TU	Per Cent
Outgoing from Vancouver and tributaries	190	42
Incoming to Vancouver	230	51
Through switched to toll centers and tributaries	33	7
Outgoing traffic is split as follows:		
Intra-area	128	67
To United States	38	20
To Canada	24	13
Through switched traffic is split as follows:		
From the area	20.2	61
From United States	4.6	14
From Canada	8.2	25

to a target date about 5 years hence.

The Canadian province of British Columbia represents one numbering plan area (604) in continent-wide toll dialing. There is one metropolitan area, Vancouver, with 32 central office units and a total of 263,000 stations. Several of the central offices will be manual for a number of years to come, with incoming toll traffic to be completed through call indicator type of equipment. Another multioffice area, Victoria, has seven central office units and 50,000 stations. There are 41 exchange areas with from 1,000 to 10,000 stations each, single office or multioffice. The total number of central offices in the province is 165; the total number of stations, 455,000; the number of toll centers, 55. Some of the toll centers and tributaries will remain manual for several years after the target date.

As shown in Fig. 1, Vancouver is the main switching center of the area both for intra-area and interarea traffic, though, naturally, direct circuits exist between many of the other toll centers. Vancouver is connected by direct circuits to all but 20 toll centers, and while these 20 toll centers are at present reached through tandem points many of them will be eliminated. There are direct circuit groups from Vancouver to five major and two other Canadian toll centers outside British Columbia. All traffic from Vancouver to the United States is switched through Seattle, Wash., with high usage groups also to Portland, Oreg., and to two smaller near-by toll centers. All local switching equipment in Canada is of the Strowger type or other step-by-step types, but two *4A* crossbar toll switching offices are planned for Montreal and Toronto.

There will be 60 toll circuit groups with about 600 circuits connected to the Vancouver toll switching center and an automatic (cordless type) toll board of about 200 positions. Table I shows the busy-hour traffic volume, with the breakdown for the outgoing and through switched traffic. The traffic volume is expressed in traffic units (TU), one TU being equal to 36 hundred-second unit calls. Table I shows the important fact that about 93 per cent of the traffic is either originated by Vancouver operators or is terminated in the Vancouver local area; therefore any equipment layout should be so planned that it will provide maximum economy in the handling of this 93 percent share of the total traffic.

Standard two-five numbers are being assigned to the stations of all central offices in the area. In Vancouver, during a transition period, mixed 6-digit and 7-digit numbers will exist, each consisting of two letters and four or five numerals. For toll dialing purpose uniform 7-digit numbers are used; in 6-digit offices the numerical equivalent of the third digit of the central office name provides the missing digit. Outside of Vancouver all newer offices are arranged for two-five numbering, while the ultimate central office codes are being assigned to older offices as well as manual offices. Such offices will be converted to standard two-five numbering as soon as practical.

The incoming traffic to Vancouver local stations will now be described. Fig. 2 shows a portion of the Strowger-type intertoll switch train. Trunks to single-unit Vancouver offices are connected to intertoll third-selector levels, and trunks to multiunit offices are connected to intertoll second-selector levels. Where separate trunk groups are provided from third-selector levels to each unit, additional codes can be made available for other central offices. Trunks to central offices having 6-digit station numbers for local dialing can be connected to the intertoll third-selector level corresponding to the third letter of the office name, or to the second-selector level, with the third digit absorbed in the toll switching trunk relay equipment. The latter practice is wasteful of central office codes because nine other codes are thus made unavailable, and it is recommended only in areas where a surplus of codes exist. However, where this practice is feasible, a rank of selectors can be eliminated.

Examining Fig. 2, it will be seen that any call, wherever it may originate, may be extended to any Vancouver station by the originating operator setting up a 7-digit directory number over the intertoll switch train and selector switches in the local office. If the call originates in any of the toll centers within the local numbering plan area, the operator reaches a trunk terminating in Vancouver and dials the 7-digit number for the direct control of the switch train in Vancouver. Calls from outside the local numbering plan area may go over direct trunks to Vancouver and the procedure is then the same as for intra-area calls; or if the call has to be switched at intermediate points outside the area, the area code 604 will direct the call to Vancouver, where the remaining seven digits will set up the switch train, as before. Vancouver toll operators may extend calls to local stations by dialing 7-digit numbers over the same switch train. Therefore one common switch train carries the traffic originating from all possible sources, and, as will be

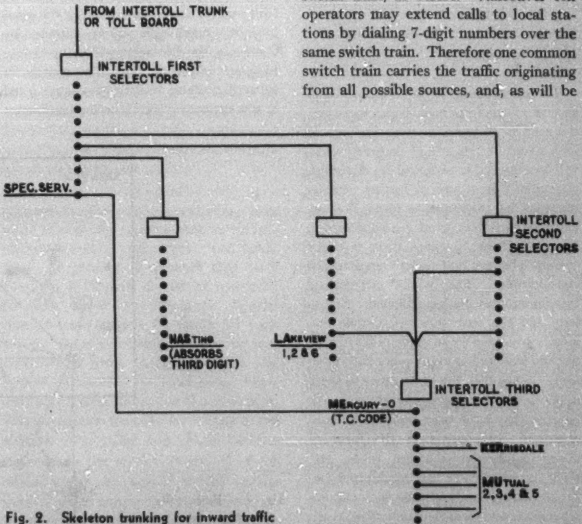


Fig. 2. Skeleton trunking for inward traffic

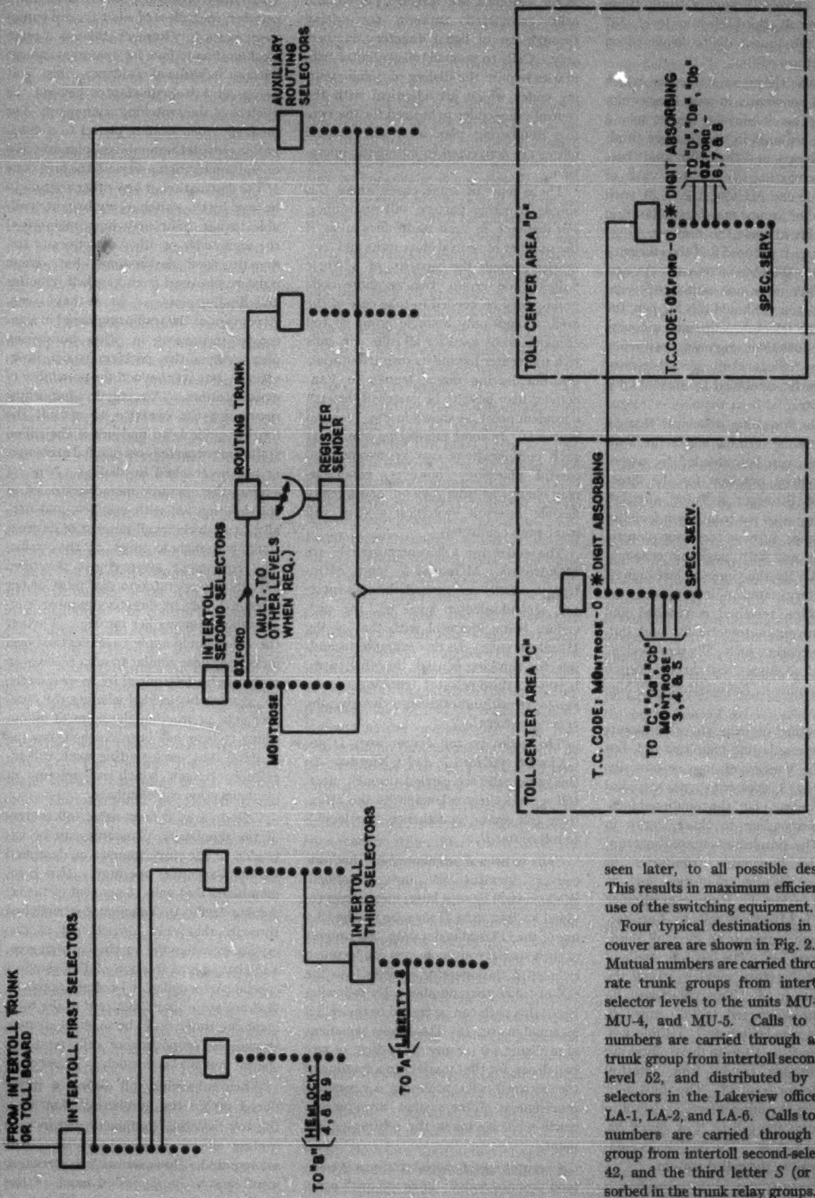

FROM INTERTOLL TRUNK OR TOLL BOARD
INTERTOLL FIRST SELECTORS
TO "B" HEMLOCK-4,8 & 9
TO "A" LIBERTY-8
INTERTOLL THIRD SELECTORS
INTERTOLL SECOND SELECTORS
OXFORD
MONTROSE
(MULT. TO OTHER LEVELS WHEN REQ.)
ROUTING TRUNK
REGISTER SENDER
AUXILIARY ROUTING SELECTORS
TOLL CENTER AREA "C"
T.C. CODE: MONTROSE-0
* DIGIT ABSORBING
TO "C", "Ca", "Cb" MONTROSE-3,4 & 5
SPEC. SERV.
TOLL CENTER AREA "D"
T.C.CODE: OXFORD-0
* DIGIT ABSORBING
TO "D", "Da", "Db" OXFORD-6,7 & 8
SPEC. SERV.

Fig. 3. Skeleton trunking for outward and through traffic

seen later, to all possible destinations. This results in maximum efficiency in the use of the switching equipment.

Four typical destinations in the Vancouver area are shown in Fig. 2. Calls to Mutual numbers are carried through separate trunk groups from intertoll third-selector levels to the units MU-2, MU-3, MU-4, and MU-5. Calls to Lakeview numbers are carried through a common trunk group from intertoll second-selector level 52, and distributed by incoming selectors in the Lakeview office to units LA-1, LA-2, and LA-6. Calls to Hastings numbers are carried through a trunk group from intertoll second-selector level 42, and the third letter *S* (or 7) is absorbed in the trunk relay groups. Calls to Kerrisdale numbers (an office which is to be retained on a manual basis for some time) are carried through a trunk group from intertoll third-selector level 537 with the subsequent digits displayed on call indicator positions.

Fig. 2 also shows an artifice by which additional economies in switching equipment may be obtained. Where only a few levels are used in two or more third-selector groups, and the codes used there are nonconflicting (as in the case of KERrisdale and MUtual 2, 3, 4, 5), such third-selector groups may be combined in one; in this manner the number of selector groups is reduced and the efficiency per selector thereby increased. As new central office units are added code conflicts may occur. Should this happen, the combination selector groups can, of course be split up into their original constituents. In smaller offices even second-selector groups can be combined by careful code assignment.

From the foregoing it is clear that all incoming traffic terminating in the Vancouver area can be extended by use of listed directory numbers for the direct pulsing of Strowger selector switches without any need for translation facilities or any other type of common control equipment, and with maximum economy and speed. By the time the last digit of the directory number is received, the called station terminal is reached and ringing current automatically sent out. As has been seen, such calls represent 87 per cent of all inward and through switch traffic incoming to the intertoll switching center.

The routing of calls to other points within the numbering plan area 604, but outside the Vancouver area, is now discussed. Fig. 3 represents again a typical skeleton routing plan, the numbered locations corresponding to those shown in Fig. 1. The simplest cases are those represented by Vancouver tributaries, or toll centers without tributaries, such as *Va* and *A* in Fig. 1. If the central office code of *A* is LIberty-8, calls to *A* are carried through trunks from intertoll third-selector level 548. The last four digits of the 7-digit directory number set up the switch train in *A*. Two tributaries *Ba* and *Bb* are connected to toll center *B*. The codes are HEmlock-4, HEmlock-8, and HEmlock-9. Calls to these offices are carried through trunks from intertoll second-selector level 43. The third digit operates the incoming intertoll selector in *B* and routes the call to *B* or *Ba* or *Bb*, as the case may be. The final digits complete the call to the called station.

In the foregoing cases the procedure is identical with that used in extending calls to Vancouver stations, i. e., Strowger selector switches are directly pulsed and calls completed without any delay through use of listed directory numbers only. Calls to manual toll centers or tributaries require the dialing of 3-digit routing codes which are identical with the central office codes to be used for the two-five numbering plan after the manual offices are converted to automatic operation.

There may be some cases where this simple trunking pattern will not suffice. For instance, in some multioffice areas, if the number of central office units and tributaries exceeds the capacity of a single 2-digit code group, two or more code groups may be needed to take care of the area, though only a single group of toll trunks is to be used for all calls; or, calls to a toll center having its own tributaries, but not having direct trunks to Vancouver, may have to be switched through a tandem point, as shown in Fig. 1 for toll center *D*. In some numbering plan areas such complications can be avoided by careful planning; however, conditions may change as time goes on, and a really flexible intertoll switching system will meet such conditions.

The codes for toll center area *C* are MOntrose-3, MOntrose-4, and MOntrose-5. Calls are carried through intertoll second-selector level 66, the procedure being identical with that of the HEmlock calls. In this example second-selector trunking is used; in other cases, however, third-selector trunking can be employed without otherwise changing the general pattern.

The codes to toll center area *D* are OXford-6, OXford-7, and OXford-8. In this case, calls are carried through intertoll second-selector level 66 to the MOntrose toll center, and thence over level 9 to toll center *D*.

Calls to area *D* originated by the Vancouver operator are now described. While, as will be seen later, means are provided so that area *D* may be reached by use of the OXford codes only, the amount of such special routing means can be much reduced if the operator prefixes 0 to the OXford directory number. By following Fig. 3 this path can be traced to the called terminal in area *D*. The simple expedient of prefixing a 0 (or any other digit, or any two digits, for that matter) to a fraction of the outward calls, while not a necessary requirement of the system, introduces a worth-while saving in the routing equipment.

Operator use of special routings is common practice today; however, with careful layout of the numbering plan, such arbitrary routings can largely be eliminated in favor of use of the listed directory number, regardless of what type of equipment is used. However, there is a practical limit as to how far operators can get without additional guidance, this limit being set by circumstances beyond the control of the switching equipment. For instance, when a call is placed to a destination outside the numbering plan area, the operator must first ascertain the area code of the destination or any other routing to be used by the switching equipment available to her. She may have memorized the area codes or other directives of frequently used destinations, but others must be obtained from keyshelf bulletins and flip-flop guides. As we have seen, 33 per cent of the traffic originated in Vancouver terminates in other numbering plan areas, so this problem is not insignificant, but it is beyond the possibility of mechanization. An equally important reason for the operator to consult the routing guide is to find out if the called station is connected to a manual exchange, or can be reached by dialing. Now, of course, the operator memorizes a lot of this information with practice, and usually a relatively small number of destinations will include most of the traffic. The equipment arrangement described permits the operator to dial most of her intra-area calls by directory number, considerably reducing her burden. Whether the same should apply to all destinations, including those similar to area *D*, must in each case be determined by its own merit; i.e., it must be decided whether the operator gets, or whether the point of diminishing returns has been reached, bearing in mind that some routing work will inevitably remain which no amount of mechanization can eliminate.

Calls to area *D* from other toll centers of the numbering plan area can be extended in the same manner as described for the Vancouver operators. If it is remembered that only 13 per cent of the incoming traffic to Vancouver is switched through, that only 61 per cent of this originates within the numbering plan area, and that only a fraction of this requires special prefix codes, it becomes apparent that not more than 2 per cent of the total incoming traffic could be so affected, even should the proportion of calls requiring special prefixes be as much as 25 per cent.

When subscriber toll dialing is introduced within the numbering plan area, the toll ticketing equipment will include routing directives where such directives are required. However, with the arrangement now to be described much of the need for such directives can be eliminated.

Through switch calls from other numbering plan areas, i.e., from the United States and from other Canadian provinces, can be set up by the direct pulsing of the 7-digit directory number into the Strowger-type intertoll switch train, except in the case of such destinations as area *D*. In such cases, operators in nearby toll centers and others having a high volume of traffic with British Columbia may receive the same routing directives as the Vancouver and other British Columbia operators; i.e., to prefix 0 before the full directory number of area *D* stations. In other switching centers having a high volume of traffic with British Columbia, foreign area translation equipment may automatically furnish the routing directives to the operator-dialed or subscriber-dialed calls to toll center *D*. However, it is probable that there will still be a certain amount of traffic to area *D* for which it will not be practical to distribute routing instructions to the many toll centers from which these calls might originate; nor will it be practical to provide automatic routing directives at all distant intertoll switching centers. Such residual traffic can be taken care of by the arrangement shown in Fig. 3.

Any toll center, distant or nearby, within or outside of the numbering plan area, deliberately or contrary to instructions, may wish to reach toll center *D* by dialing, say OXford-3. A group of so-called "routing" trunks are connected to level 69 of the intertoll second selectors. The purpose of these trunks is to seize a simple register-sender, and to indicate to the latter what routing digits are to be sent out. The routing trunks are very simple devices, consisting of a few relays only. Since the traffic they receive from a level is obviously very small, as shown in the previous discussion, each trunk group therefore should preferably serve several trunk routes similar to those serving area *D*, and the proper routing directive can be given by the routing trunk to the register-sender corresponding to the level over which it was seized. This register-sender is the same as the one used for simplified automatic alternate routing.[3] It receives and stores incoming dial pulses; but even before that it sends out the routing digits (69 in example for calls to toll center *D*) as prescribed by the routing trunk, which are then followed by the digits stored in the register-sender. A Strowger-type auxiliary routing selector is associated with each routing trunk, having those toll trunks connected to its various levels over which the calls received by the routing trunks are to be extended. For instance, the trunks to toll center *C* are connected to level 6 of the auxiliary routing selector, and calls to area *D* are first routed to these trunks over level 6 and then over level 9 of the incoming selector in toll center *C*.

This is all that is required to provide toll dialing by the use of directory numbers throughout. The equipment needed in a fairly large intertoll switching center, such as Vancouver, is now described. To be on the most liberal side, it will be assumed that all through switch traffic, originating in British Columbia or at any other place outside, should be extended by the dialing of two-five directory numbers, and that only 75 per cent of the traffic can be handled without additional digits, such as areas *A*, *B*, *C*, or *Va*, while 25 per cent is of the type exemplified by area *D*. The total through switch traffic is 33 TU, 25 per cent of which is 8.25 TU. Two groups of routing trunks will handle such traffic, with 4.125 TU per group. At the grade of service 1 in 1,000, each group requires 13 trunks, or 26 for the whole office plus 26 auxiliary routing selectors. If register-senders are already provided for alternate routing purposes, one or two additional register-senders would be needed for the routing of these calls. The routing trunks have 20 relays each and a register-finder rotary switch. The intertoll selectors are standard Strowger switches with four relays and 660-point banks. However, it appears to be a more practical assumption that for all through switch traffic originating in the numbering plan area, and for half the traffic originating outside, the operators or other machines will do the routing, so that only 25 per cent of 6.4 TU, i.e., 1.6 TU, will pass through routing trunks. The number of routing trunks and selectors can now be reduced to ten, and most likely no register-senders other than those used for alternate routing will be needed.

What has been achieved is now reviewed here. All traffic incoming into or switched within a whole numbering plan area is extended by the dialing of the two-five directory numbers. All but 0.6 per cent or, at the least, all but 3 per cent of this traffic will be set up by straightforward pulsing methods on inexpensive direct pulse-controlled Strowger selectors, and the rest of the traffic channeled through special routing equipment, which is negligibly small in view of the size of the installation. Actually, the inward and through traffic is only 58 per cent of the total; therefore the fraction of specially routed traffic is only about half of the figures quoted here. Furthermore, it is axiomatic that the highest speed of service can be achieved only when the switch train is directly set up by the incoming pulses, while storage, translation, and retransmission must necessarily introduce some delay. It is not only that costly and complex equipment, as needed in systems with 100-per-cent common control, is eliminated for 98 per cent or more of the traffic but also that the highest possible speed is retained for practically all the traffic. Because outpulsing commences from the routing trunks as soon as they are seized, there is no added delay on calls using these trunks. Furthermore, it is obvious that the routing trunks may be added, removed, or rearranged where and when needed, with no interference with the rest of the equipment, and without affecting the operation of the system. Therefore the basic flexibility of the Strowger system is retained; in this case it is exemplified by the fact that common control equipment need be provided only when it does some good, and not because the system is inoperative without it.

Next the routing of calls to destinations outside the numbering plan area is to be considered. As has been seen, all such traffic will be extended through six principal outlets (five in Canada and one in the United States) and five other toll centers (two in Canada and three in the United States). Each of these 11 points is reached by the dialing of a 2-digit followed by a 10-digit or 7-digit directory number, or by any other special procedure required by the equipment at those distant points. These trunks are connected to vacant intertoll second-selector levels, preferably 1 and 0. The use of such prefix codes can be considered as the automatic cordless toll board equivalent of plugging into the outgoing jack multiple of the trunk groups to those destinations on conventional cord-type toll boards, or the routing codes used in toll tandem practice.

Toll center codes may be assigned as needed. As previously mentioned, for smaller toll centers the central office code serves also as a toll center code. In Vancouver, as shown in Fig. 2, special operators and plant test facilities can be reached through a selector train connected to the first level of the intertoll first selectors, without any special toll center code. However, to accommodate toll switching equipment at distant control switching points, which requires distinct toll center codes, MErcury-0 may be considered as the toll center code, and level 1 of the intertoll first selectors and level 630 of the intertoll third selectors are tied together for that purpose. MOntrose-0 and OXford-0 were assigned as toll center codes for *C* and *D* respectively, as seen in Fig. 3.

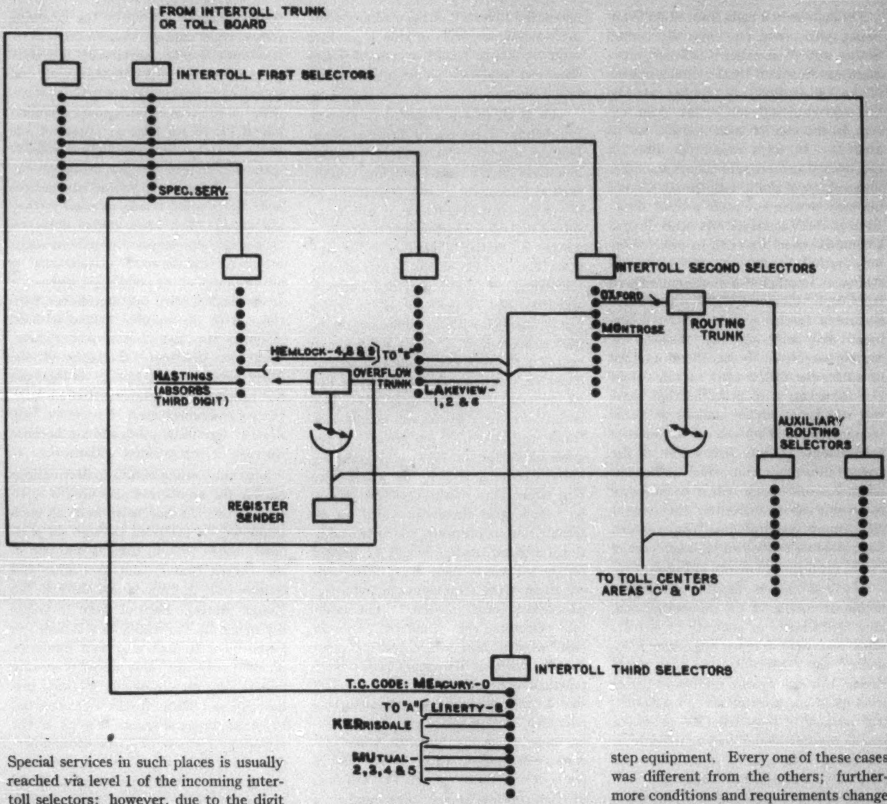

FROM INTERTOLL TRUNK
OR TOLL BOARD
INTERTOLL FIRST SELECTORS
SPEC. SERV.
INTERTOLL SECOND SELECTORS
OXFORD
MONTROSE
ROUTING
TRUNK
HEMLOCK-4,8 & 9
TO "B"
HASTINGS
(ABSORBS
THIRD DIGIT)
OVERFLOW
TRUNK
LAKEVIEW-
1,2 & 6
AUXILIARY
ROUTING
SELECTORS
REGISTER
SENDER
TO TOLL CENTERS
AREAS "C" & "D"
INTERTOLL THIRD SELECTORS
T.C. CODE: MERCURY-0
TO "A"
LIBERTY-8
KERRISDALE
MUTUAL-
2,3,4 & 5

Fig. 4. Combined skeleton trunking

Special services in such places is usually reached via level 1 of the incoming intertoll selectors; however, due to the digit absorbing feature on their 0 level, it is immaterial whether the third digit 0 is dialed or not.

The topography and other characteristics of toll centers in British Columbia limits the extent to which alternate routing of toll calls is feasible. By the visual display of group busy conditions to each of the Vancouver toll operators, the latter are readily able to handle the alternate routings themselves, in fact, as mentioned earlier in this paper, with a greater degree of flexibility than by any automatic means. There are various indications to guide the distant operator if congestion is encountered in the process of switching a through call, and simple automatic alternate routing equipment[3] may be added where and when needed. Here again, through switch traffic is only 7 per cent of the total, and only a fraction of this over routes where alternate routing is feasible; therefore only the simplest automatic alternate routing equipment can be justified, if any at all.

Fig. 4 combines the arrangements shown separately in Figs. 2 and 3, as well as some of the additional provisions previously described. Of course, Fig. 4 represents only a fraction of the inter-toll switch train at Vancouver, and is shown here only as a typical example of the general principles which have been discussed.

The author had the opportunity for a number of years to be actively engaged in planning several large toll dialing areas to be integrated in the continental network. He found that the most economical as well as the most suitable arrangement could be assembled for each particular application owing to the flexibility of a kind of building-block technique, with step-by-step equipment. Every one of these cases was different from the others; furthermore conditions and requirements change all the time. Many of these conditions and requirements are unforeseeable between the time when the equipment is initially planned and the time when it is placed in service. Usually during the initial service period the conditions are relatively straightforward, and with the technique described in this paper the initial investment is at a minimum, yet the system provides the highest speed of operation. As time passes, new and previously unpredictable conditions arise. Equipment specially designed to take care of these conditions can be added, but because usually only a fraction of the traffic is affected the added equipment is concentrated only at those points where it is actually needed, which is obviously also sound economic policy.

The same principles were embodied in the *FW-1* intertoll switching system recently developed by the Automatic Electric Company for continent-wide toll dialing, described in this paper, as well as in the company's four automatic toll ticketing systems. The features of the *FW-1* intertoll switching system include 4-wire switching of transmission paths throughout; pad switching; pulsing and supervisory signals completely separated from the transmission paths; silver inlaid bank contacts; integrated high-speed automatic cordless toll board with a new, efficient method of automatic call distribution and new, inexpensive keysenders; and automatic alternate routing and code conversion, as discussed in this paper. It can be planned without any common control, or with partial common control, or with full common control, whether introduced initially or added at any time later and at any point; automatic camping or other disposition of calls under delay and overflow conditions; a new method of centralized automatic traffic recording of the whole system; centralized control and supervision of traffic operation; high efficiency service observation; plant test test facilities; self-contained power and supervisory equipment; etc. The equipment for the first application of this system is being manufactured for New Westminster, British Columbia, and is to be placed in service in 1955.

References

1. TRANSMISSION DESIGN OF INTERTOLL TELEPHONE TRUNKS, H. R. Huntley. *AIEE Transactions*, vol. 72, pt. 1, Nov. 1953, pp. 670–76.

2. TRAFFIC ENGINEERING TECHNIQUES FOR DETERMINING TRUNK REQUIREMENTS IN ALTERNATE ROUTING TRUNK NETWORKS, C. J. Truitt. *Bell System Technical Journal*, New York, N. Y., March 1954, pp. 277–302.

3. COMMON CONTROL AND THE STROWGER SYSTEM, C. F. Ffolliott. *Automatic Electric Technical Journal*, Chicago, Ill., March 1954, p. 120.

A Rectifier Algebra

DAVID H. SCHAEFER
ASSOCIATE MEMBER AIEE

THIS paper presents an algebra that can aid in the simplification, the synthesis, and the understanding of circuits containing rectifiers and resistors. More specifically, it gives a method of rendering rectifier-resistor circuits into symbols with rules for manipulating these symbols. Here, in essence, an algebra is set up analogous to, but differing considerably from the well-known algebra of Shannon for relay circuits.[1]

This is the initial result of a project the final aim of which is the development of an algebra suitable for the simplification and analysis of magnetic amplifier circuits. The investigation has required the study of symbolic logic and its present applications. Shannon's switching algebra is the Boolean algebra of classical 2-valued symbolic logic. This can deal with only two values, in logic called "true" and "false," in relay circuits "1" and "0." Besides the classical 2-valued system, there are other systems of logic, most of which deal with three or more values. The motivation for the algebra presented here is the infinite-valued logic of Lukasiewicz and Tarski.[2–4] (Reference 4 gives a discussion in English of the three- and *m*-valued cases of this logic.) Most of the circuits presented in this paper can be expressed completely within the Lukasiewicz-Tarski logic. It was found, however, that the simpler algebra presented here could be constructed with the use of two symbols from the Lukasiewicz-Tarski logic along with plus and minus signs which are not part of the Lukasiewicz-Tarski system.

In the next section, the details of this rectifier algebra will be given and in the last section a few simple examples of its use will be shown. No knowledge of symbolic logic or switching algebra is required for an understanding of the following.

The Rectifier Algebra

The algebra to be used here is not the ordinary numerical algebra taught in high school, nor is it the Boolean algebra used for relay circuits. It is, rather, something of a union of the two.

Let a, b, c, etc., be any real numbers, positive or negative. Two new symbols, $\vee$ and $\wedge$, may be defined as

$a\vee b$ (read a max b) = max $[a, b]$, i.e., $a\vee b$ has the value of the more positive of a or b

$a\wedge b$ (read a min b) = min $[a, b]$, i.e., has the value of the more negative of a or b

It can be seen that in essence only one new symbol has been introduced as $\vee$ and $\wedge$ can be defined in terms of each other and a minus sign, as follows

$$a\vee b=-(-a.\wedge.-b) \tag{1A}$$

where $-a.\wedge.-b=(-a)\wedge(-b)$, the dots replacing parentheses. Also

$$a\wedge b=-(-a.\vee.-b) \tag{1B}$$

Note that parentheses are not needed for a line of $\vee$'s or $\wedge$'s, i.e

$$(a\vee b)\vee c=a\vee(b\vee c)=a\vee b\vee c \tag{2}$$

$$(a\wedge b)\wedge c=a\wedge(b\wedge c)=a\wedge b\wedge c \tag{3}$$

There are two obvious simplification rules (identical with those of the 2-valued logic)

$$a\vee a=a \tag{4}$$

$$a\wedge a=a \tag{5}$$

The following governs the movement of a minus sign through a parenthesis

$$-(a\vee b)=-a.\wedge.-b \tag{6}$$

$$-(a\wedge b)=-a.\vee.-b \tag{7}$$

Note that the $\vee$ sign becomes a $\wedge$ sign and vice versa in this operation.

The plus sign, which retains its conventional algebraic meaning, combines with the minimum and maximum operations as follows.

$$a+(b\vee c)=(a+b)\vee(a+c) \tag{8}$$

$$a+(b\wedge c)=(a+b)\wedge(a+c) \tag{9}$$

Lastly, the rules for combining $\vee$'s and $\wedge$'s are

$$a\vee(b\wedge c)=(a\wedge b)\wedge(a\vee c) \tag{10}$$

$$a\wedge(b\wedge c)=(a\vee b)\vee(a\wedge c) \tag{11}$$

Some Rectifier Circuits and Their Equations

It will now be shown by example how the algebra just presented can be applied to certain rectifier-resistor circuits, using either a-c or d-c sources.

Fig. 1 shows some simple circuits and their equations. The reader may examine each circuit and equation, and convince himself that the equations are correct. The equation of Fig. 1(A), for instance, says that voltage e_0 is equal to voltage a if a is more positive than

---

Paper 54-516, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted October 19, 1953; made available for printing August 30, 1954.

DAVID H. SCHAEFER is with the Naval Research Laboratory, Washington, D. C.

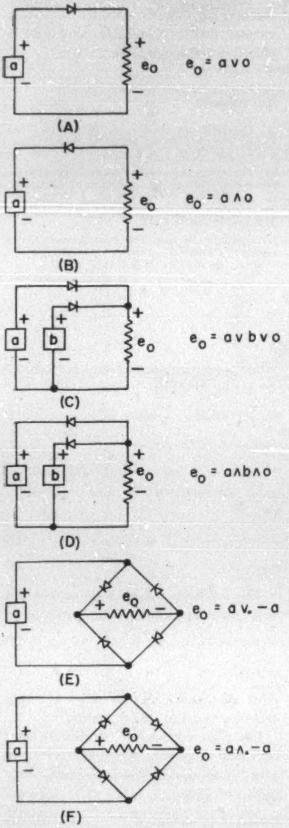

+
a
−
+
e_0
−
$e_0 = a\vee 0$
(A)
$e_0 = a\wedge 0$
(B)
b
$e_0 = a\vee b\vee 0$
(C)
$e_0 = a\wedge b\wedge 0$
(D)
$e_0 = a\vee.-a$
(E)
$e_0 = a\wedge.-a$
(F)

Fig. 1 (left). Some basic rectifier-resistor circuits and their equations

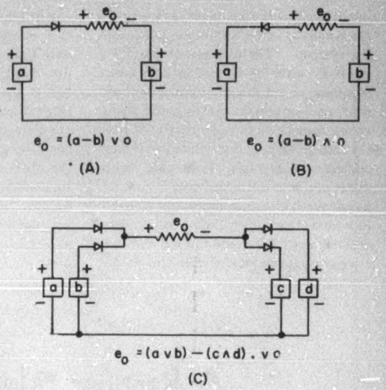

+ e_0 −
a
b
$e_0 = (a-b)\vee 0$
(A)
$e_0 = (a-b)\wedge 0$
(B)
c
d
$e_0 = (a\vee b)-(c\wedge d).\vee 0$
(C)

Fig. 2 (right). Diagrams and equations for circuits containing several voltage sources

zero, and that e_0 is zero if voltage a is more negative than zero. If the circuits in pairs are examined, a and b, c and d, and e and f, it is seen that reversing rectifiers in the diagrams is equivalent to changing v's to ∧'s and ∧'s to v's.

It is obvious, without the use of any algebra, that Fig. 1(F) is Fig. 1(E) with the positive polarity of e_0 considered to be at the opposite end of the resistor. For the purpose of becoming more familiar with the algebra, it is shown that the equations agree with this self-evident fact. To prove e_0 of Fig. 1(F) $= -e_0$ of Fig. 1(E)

$-e_0$ of Fig. 1(E) $= -(a\vee.-a)$
$= -a.\wedge.-(-a)$ by equation 6
$= -a.\wedge a$
$= e_0$ of Fig. 1(F)

Fig. 2 shows diagrams and equations for circuits containing several voltage sources. Fig. 3 presents a circuit where a rectifier is in parallel with a resistor.

A demonstration can now be given of how the algebra can assist in the simplification of circuits. First, the circuit of Fig. 4 will be simplified. This is a combination of the circuits of Fig. 1(E) and Fig. 2(A).

Here

$e_1 = a\vee.-a$

and

$e_0 = (e_1-a)\vee 0$
$e_0 = (a\vee.-a)-a.\vee 0$ by substitution
$= a-a.\vee.-a-a.\vee 0$ by equation 8
$= 0\vee.-2a.\vee 0$
$= -2a.\vee 0$ by equation 4

The circuit for this equation is Fig. 1(A), where the voltage source equals $-2a$, and is shown as Fig. 5. This simplification could have been arrived at without the use of the algebra, but it appears that even for as simple a circuit as this, the algebra speeds the solution of the problem.

Fig. 6 is a slightly more complicated circuit. Here

$e_1 = (a-b)\vee 0$
$e_2 = (c-b)\vee 0$
$e_0 = e_1\vee e_2\vee 0$
$e_0 = (a-b)\vee 0.\vee.(c-b)\vee 0.\vee.0$
$= (a-b)\vee 0\vee(c-b)\vee 0\vee 0$ by equation 2
$= (a-b)\vee(c-b)\vee 0$ by equation 4
$= (a\vee c)-b.\vee 0$ by equation 8

This is a special case of Fig. 2(C) and the simplified circuit is illustrated as Fig. 7.

Finally, a different type of problem is shown. Fig. 8(A) is a common circuit for obtaining the lesser of two voltages.

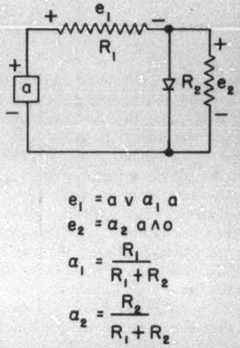

+ e_1 −
R_1
+
a
−
R_2
+
e_2
−
$e_1 = a\vee a_1 a$
$e_2 = a_2 a\wedge 0$
$a_1 = \frac{R_1}{R_1+R_2}$
$a_2 = \frac{R_2}{R_1+R_2}$

Fig. 3. Circuit with rectifier in parallel with a resistor

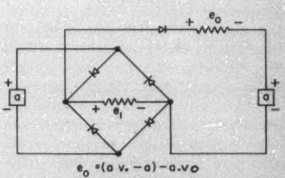

+ e_0 −
+
a
−
+ e_1 −
$e_0 = (a\vee.-a)-a.\vee 0$

Fig. 4. A combination of the circuits of Figs. 1(E) and 2(A)

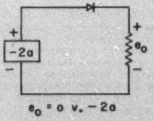

+
−2a
−
+
e_0
−
$e_0 = 0\vee.-2a$

Fig. 5. Simplified version of Fig. 4

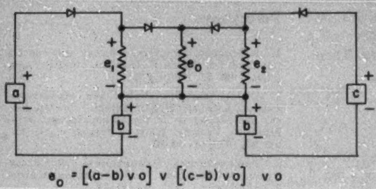


Fig. 6 (left). A combination of the circuits of Figs. 2(A) and 1(C)

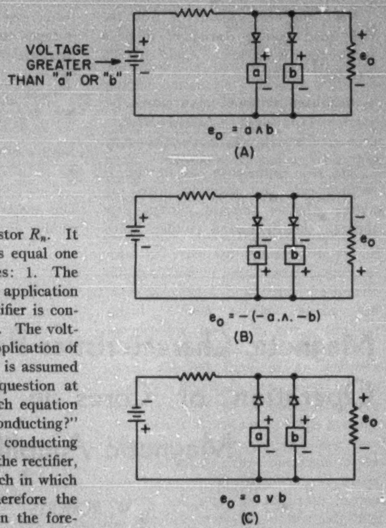


Fig. 8. A—Common circuit for obtaining lesser of two voltages. B—Circuit for obtaining greater of two voltages. C—Redrawn version of B

It should be noted that the equation for this circuit is $a \wedge b$ in contrast to $a \wedge b \wedge 0$ for Fig. 1(D). A practical problem might be as follows: Given the circuit of Fig. 8(A), find a circuit that will select the greater of two voltages. In symbols, a circuit for $a \wedge b$ is given, and a circuit for $a \vee b$ is wanted. The solution follows.

Note that $a \vee b = -(-a.\wedge.-b)$ by equation 1(A). The circuit for $-a.\wedge.-b$ is that for $a \wedge b$ with the polarity of the voltage sources a and b reversed. Therefore, the circuit for $a \vee b$ is that for $a \wedge b$ with the polarities of a and b reversed and the opposite side of the output resistor considered positive. This is shown as Fig. 8(B) and redrawn as Fig. 8(C).

The three simple examples of the use of the algebra should be indicative of its possible use with more complicated circuits.

It was stated at the start that these are the very early results of a project whose aim is the evolution of an algebra capable of handling magnetic amplifier circuits. Magnetic amplifiers generally contain three elements: magnetic cores, rectifiers, and resistors. Here is demonstrated an algebra that can deal with two of the three.

A General Method

A general method has been evolved whereby equations can be written for any circuit containing rectifiers, resistors, and voltage sources. Consider first a complicated network that contains voltage sources, resistors, and only one rectifier. The problem is to solve for e_n the voltage across the resistor R_n. It is known that e_n will always equal one of the two following voltages: 1. The voltage obtained for e_n from application of Kirchoff's laws if the rectifier is considered an open circuit; or 2. The voltage obtained for e_n by the application of the same laws if the rectifier is assumed a short circuit. The only question at any particular time is "Which equation is valid?" or "Is the rectifier conducting?" At the boundary between conducting and nonconducting states of the rectifier, no voltage is across the branch in which the rectifier is located. Therefore the two solutions for e_n noted in the foregoing are equal since the resistive value assumed for this branch cannot affect the value of e_n at this crossover point. There are two solutions and we wish to transfer from one to the other when they are equal. This can be accomplished by putting the sign $\vee$ or $\wedge$ between the two solutions. Which sign to so place is determined by noting whether current flow through the rectifier in its forward direction tends to increase or decrease e_n. If such a current flow tends to increase e_n, a $\vee$ sign is placed between the solutions. If such a current tends to decrease e_n, a $\wedge$ sign is placed between the solutions.

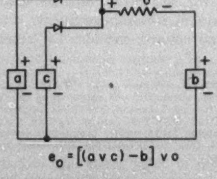


Fig. 7. Simplified version of Fig. 6

The circuit of Fig. 3 is a very simple example. If the rectifier is open, $e_1 = \alpha_1 a$ and $e_2 = \alpha_2 a$. If the rectifier is a short circuit, $e_1 = a$ and $e_2 = 0$. Forward current flow through the rectifier tends to increase e_1 and decrease e_2. Therefore $e_1 = a \vee \alpha_1 a$ and $e_2 = \alpha_2 a \wedge 0$.

In a circuit which has many rectifiers, the solution requires a series of calculations of the type just described. Solutions for all possible combinations of rectifiers open and closed must be written. For N rectifiers there are $2N$ such solutions. These solutions are linked by $\vee$ and $\wedge$ signs determined by noting whether forward current through successive rectifiers increases or decreases the desired voltage.

As an example of the use of the method for a 2-rectifier circuit, we will solve for e_0 of Fig. 9. There are four possible modes of operation for the circuit

$$e_0 = \frac{aR_1R_2 + bR_3R_2}{R_1R_3 + R_1R_2 + R_2R_3} \text{ when both rectifiers are conducting}$$

$$e_0 = \frac{aR_2}{R_2 + R_3} \text{ when only the right hand rectifier is conducting}$$

$$e_0 = \frac{bR_2}{R_1 + R_2} \text{ when only the left hand rectifier is conducting}$$

$e_0 = 0$ when neither rectifier is conducting

First, the expression for e_0 with the right-hand rectifier considered to be an open circuit is written. This is a 1-rectifier circuit, so the method just demonstrated can be utilized.

$$e_0 \text{ (right-hand rectifier open)} = 0 \wedge \frac{bR_2}{R_1 + R_2}$$

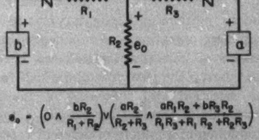

R₁
R₃
R₂
e₀
b
a
e₀ = (0 ∧ bR₂/(R₁+R₂)) v (aR₂/(R₂+R₃) ∧ (aR₁R₂ + bR₃R₂)/(R₁R₃+R₁R₂+R₂R₃))

Fig. 9. A circuit with four modes of operation

Next, the expression for e_0 with the right-hand rectifier considered to be a short circuit is written. Again this is a 1-rectifier circuit.

$$e_0 \text{ (right-hand rectifier short-circuited)} = \frac{aR_2}{R_2+R_3} \wedge \frac{aR_1R_2+bR_3R_2}{R_1R_3+R_1R_2+R_2R_3}$$

The two expressions, as in the 1-rectifier case, are similarly linked by noting whether forward current flow through the right-hand rectifier tends to increase or decrease e_0. In this case it tends to increase e_0 no matter what the state of the left-hand rectifier. Therefore

$$e_0 = \left(0 \wedge \frac{bR_3}{R_1+R_3}\right) \vee \left(\frac{aR_2}{R_2+R_3} \wedge \frac{aR_1R_2+bR_3R_2}{R_1R_3+R_1R_2+R_2R_3}\right)$$

This method can be extended to any number of rectifiers.

References

1. A Symbolic Analysis of Relay and Switching Circuits, **Claude E. Shannon.** *AIEE Transactions*, vol. 57, 1938, p. 713.

2. Untersuchungen uber den Aussangenkalkul, **J. Lukasiewicz, A. Tarski.** *Comptes Rendus des Séances de la Société des Sciences et des Lettres de Varsovie*, Warsaw, Poland, classe III, vol. 23, 1930, pp. 30–50.

3. Philospshische Bemerkungen zu Mehrwentigen Systemen des Aussagenkalkuls, **J. Lukasiewicz.** *Ibid.*, pp. 51–77.

4. Symbolic Logic (book), **C. I. Lewis, C. H. Langford.** Appleton-Century Company, New York, N. Y., 1932, chap. VII.

Magnetic Characteristics Pertinent to the Operation of Cores in Self-Saturating Magnetic Amplifiers

R. W. ROBERTS
ASSOCIATE MEMBER AIEE

A METHOD of measuring the magnetic characteristics which control the manner in which cores operate in self-saturating magnetic amplifiers is necessary for both economical production and accurate design of such amplifiers. It has been found[1,2] that there is a lack of direct correlation between amplifier characteristics, such as gain, and either d-c or a-c major hysteresis loop measurements. The difficulty lies in the manner in which the time lag[3] occurring between the application of a magnetizing force on a core and the resultant change in magnetic induction, affects the operation of a self-saturating amplifier. The result is that, to determine the effective magnetic characteristics of a core pertinent to self-saturating magnetic-amplifier operation, it is necessary to measure characteristics under simulated amplifier operating conditions in such a way that the effect of magnetization lag will be similar.

A method has been described by Conrath[4] which will allow a simple and rapid measurement of core behavior under conditions approximating magnetic amplifier operation. An independent version of this test method has been used during the past year and a half for the selection and matching of many thousands of toroidal cores used in the production of a variety of different types of self-saturating magnetic amplifiers. Experience has shown that, although not perfectly accurate in every case, the test measurements will give excellent prediction of amplifier performance in a great many applications. In addition to its usefulness as a production test instrument, the test method has proved most valuable in development studies of magnetic material in the improvement of core fabrication techniques, and in the investigation of environmental effects on magnetic amplifier cores.

Introduction of this core test method has also yielded a new approach to magnetic-amplifier analysis and circuit design procedure. The core characteristics, as measured by the test method, can be used to obtain an analytical expression for the control characteristics of an amplifier. The use of this procedure has been found to be most useful in amplifier design work and particularly in the selection of appropriate cores for particular applications. Included in the scope of this paper will be a description of the core test method and its relation to magnetic-amplifier operation. In addition, the procedure for utilizing core tester measurements in the analysis or design of self-saturating magnetic amplifiers will be determined.

Core Tester Operation

The basic circuit of the core tester is illustrated in Fig. 1. It is essentially a half-wave self-saturating magnetic amplifier except for the addition of a pickup winding and the relative proportioning of the excitation or load winding and a-c supply voltage. The circuit design is such that the excitation supply applies half-wave pulses of magnetizing force to a core, of sufficient magnitude to drive the flux to positive saturation. Then, for a period of time exactly equal to 1 half-cycle of the test frequency, the excitation winding is effectively isolated by the blocking of the excitation rectifier. During this period, a d-c control demagnetizing force will be the only external magnetizing force acting on the core and, since the impedances in the control and pickup winding circuits are made to be effectively infinite, there will be negligible external loading on the core. Thus, each cycle the core flux will be driven to positive saturation and then reset under the influence of a constant demagnetizing force for a period of exactly 1 half-cycle of the test frequency. Hence the phrase "constant current flux reset core tester," has been used to describe this test method.

The total flux change each cycle, as a function of the control demagnetizing force, is required to determine the performance of a core in a magnetic amplifier. In the core tester the cyclic flux change will have small effect on the excitation current. This is particularly true when a small core is being tested with only a 1- or 2-turn excitation winding. However, if a separate pickup winding is used, the flux change may be obtained by appropriate measurement of the induced voltage, and also the measure-

---

Paper 54-392, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 15, 1954; made available for printing August 25, 1954.

R. W. Roberts is with the Westinghouse Electric Corporation, Pittsburgh, Pa.

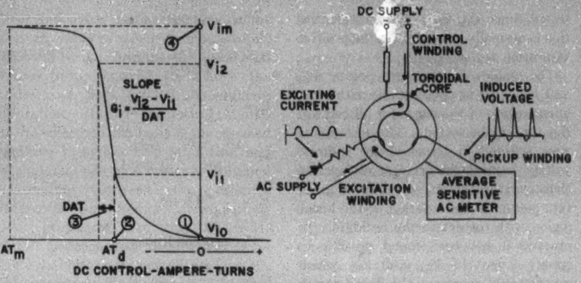


Fig. 1. Magnetic amplifier core tester: basic circuit and typical core test characteristic curve. Standard measurement points—1. V_i measured at zero control ampere-turns. 2. Control ampere-turns measured at specified voltage $V_{i,1}$. 3. Gain on linear region between specified voltages $V_{i,2}$ and $V_{i,1}$, as calculated from measured DAT. 4. $V_{i,m}$ measured at specified control ampere-turns AT_m

ment will be independent of the magnitude of excitation current so long as the core is driven to saturation each cycle.

The a-c voltmeter reading of the induced voltage in the pickup winding will be related to the cyclic flux change as follows. The instantaneous value of induced voltage per turn v_i is given by

$$v_i=(d\phi/dt)\times 10^{-8} \tag{1}$$

During a single cycle of steady-state operation the flux will change monotonic from the reset point ϕ_0 to positive saturation ϕ_m and will then return monotonic to the reset point. If the period of flux increase is from $t=0$ to $t=t_1$ and the period of flux decrease from $t=t_1$ to $t=T$, then the integral of the absolute value of v_i during each of these periods will be, from integration of equation 1

$$\int_0^{t_1}|v_i|dt=\int_{t_1}^{T}|v_i|dt=(\phi_m-\phi_0)\times 10^{-8} \tag{2}$$

Although the two periods may not be of exactly equal length, the integrals will be equal since the net flux changes must be the same.

An average sensitive a-c voltmeter will give an indication proportional to the integral over a cycle of the absolute value of an a-c voltage. Therefore, it may be used to measure the total flux change per cycle. Average sensitive a-c voltmeters are conventionally calibrated in terms of the rms value of a sinusoidal input. With this calibration, the meter reading V will be related to the integral of the instantaneous voltage v by

$$V_i=\frac{\pi}{\sqrt{2}}\frac{1}{T}\int_0^T|v|dt \tag{3}$$

Using an average sensitive a-c meter to read v_i, the meter value of induced voltage per turn V_i will be

$$V_i=\frac{\pi}{\sqrt{2}}\frac{1}{T}(\phi_m-\phi_0)\times 10^{-8} \tag{4}$$

Simplifying and using $\Delta\phi=\phi_m-\phi_0$

$$V_i=2.22f\Delta\phi\times 10^{-8} \tag{5}$$

A typical curve of induced voltage V_i versus d-c control ampere-turns AT for an oriented 50 per cent nickel-iron core is shown in Fig. 1. This curve is similar in form to the left flank of a hysteresis loop plotted in reverse with positive saturation coincident with the horizontal axis. That is, zero-induced voltage corresponds to the flux remaining at positive saturation throughout each cycle, while the maximum or saturation value of induced voltage is obtained when the flux swings from positive to negative saturation and back again each cycle. Core tester measurements may be expressed either in terms of induced voltage and control ampere-turns as in Fig. 1, or in terms of magnetic quantities such as flux density and magnetizing force. If converted into magnetic quantities, the data will represent the left flank of an effective major a-c hysteresis loop formed by the locus of the lower left-hand corners or reset points of the minor loops actually followed by the core during each cycle.

Consideration of the operation of the core tester at zero frequency or direct current will serve to illustrate further the significance of the test method. For d-c operation the test circuit will be the same except for the replacement of the

Nomenclature

A_c = effective core cross-section area, square centimeters
AT = average ampere-turns applied to core during reset period
AT_d = value of control ampere-turns required in core tester to obtain $V_i=V_{i,1}$
AT_M = value of control ampere-turns used in core tester to measure V_{im}
AT_p = peak value of excitation ampere-turns on core tester
B_m = saturation value of flux density in core, gausses
B_r = residual value of flux density in core, gausses
ΔB_1 = peak-to-peak flux density change, gausses corresponding to induced voltage per turn $V_{i,1}$
ΔB_2 = peak-to-peak flux density change, gausses corresponding to induced voltage per turn $V_{i,2}$
C = constant
DAT = difference in values of control ampere-turns required in core tester to change V_i from $V_{i},1$ to $V_{i},2$
f = a-c supply frequency, cycles per second (cps)
G_i = slope of linear region of core test characteristic curve, V/T per AT
H_d = value of magnetizing force, oersteds corresponding to AT_d
H_p = peak magnetizing force, oersteds corresponding to AT_p
H_m = value of magnetizing force, oersteds corresponding to AT_m
I_c = average d-c control current, amperes
K = gain of amplifier in terms of load voltage per control ampere-turn
l_m = mean length of magnetic path in core, centimeters
N_c = number of turns in control winding on core
N_G = number of turns in gate winding on core
R_c = control circuit resistance, ohms
R_L = load circuit resistance, ohms
T = period of a-c supply frequency, seconds
V_{ac} = a-c supply voltage, rms volts
V_G = average sensitive a-c voltmeter value of induced voltage in gate winding*
v_G = instantaneous value of voltage induced in gate winding
V_i = average sensitive a-c voltmeter value of induced voltage per turn*
v_i = instantaneous value of voltage induced per turn in a winding on the core
$V_{i,1}$ = value of induced voltage per turn used for measurement of point 2 on core test characteristic curve
$V_{i,2}$ = value of induced voltage per turn used for measurement of point 3 on core test characteristic curve
$V_{i,m}$ = maximum or saturation value of induced voltage per turn obtained from core tester when control ampere-turns equal AT_m
$V_{i,0}$ = value of induced voltage per turn obtained from core tester under conditions of zero control ampere-turns
V_L = amplifier output voltage average value for d-c output or average sensitive a-c voltmeter value for a-c output*
μ_d = value of effective differential permeability corresponding to core test characteristic gain, G_i
ϕ_m = saturation flux level in core
ϕ_0 = flux in core at end of reset period
ϕ_r = residual flux in core

---

* Average sensitive a-c voltmeter assumed by convention to be calibrated in terms of rms value of sinusoidal voltage.

average sensitive a-c voltmeter with a ballistic galvanometer or flux-meter and the use of a switch and d-c power supply for the excitation circuit. Under these conditions, when the excitation current is turned off, the flux will be reset by the control current to a point on the major d-c hysteresis loop. A plot of the data for various values of control magnetizing force will give the actual left flank of the major d-c hysteresis loop for the core.

When the tester is operated from a-c excitation, the major difference from d-c operation will be that for a certain range of control magnetizing force the flux will not reset to a point on the d-c hysteresis loop when the excitation current is blocked or turned off for a half-cycle of the a-c test frequency. This is because an appreciable length of time is required for the magnetization to reach its final state after the application of a magnetic field.[3,5] This time lag is variable with the magnitude of the applied field and the initial state of the core but it is often considerably longer than a half-cycle period of even 60-cps power frequency.[3]

The cause of this time lag is not as yet completely understood. Although eddy currents will certainly cause a time lag, the observed delays in rectangular loop materials are much greater than predicted by the usual eddy current calculations. It has been suggested that there may be a magnetic effect analogous to viscosity.[5] However recent studies of domain structures in these materials indicate that the lag is most likely attributable the eddy currents associated with the movement of the sharp change in magnetization at the domain wall.

Core Test Characteristic Values

For greatest usefulness, the important features of the core test characteristic curve must be expressed by a set of characteristic values. The standard test values described in Fig. 1 have been found convenient to measure under production conditions and, also, accurately to represent features of the core characteristic with respect to magnetic-amplifier application.

Measurement of these standard test values is dependent on several test conditions which must be specified on the basis of the characteristics of the core material being tested. First, the peak value of the excitation ampere-turns AT_p must be sufficiently great to assure saturation of the core. This value can be set on the basis of the d-c magnetic properties of the material being tested, since the time lag of magnetization is normally very small in the positive saturation region.

The values of the test points $V_{i,1}$ and $V_{i,2}$ must be chosen to lie within the linear region of the curve for all cores of the type to be tested. The test points were selected in terms of the induced voltage because the greatest variance of characteristics among cores of a given type occurs in the location of the linear region with respect to the ordinate. In practice it has been found possible to specify values of $V_{i,1}$ and $V_{i,2}$ which satisfactorily represent the linear region of the curve for all undamaged cores of any given size and material.

Measurement of the saturation induced voltage $V_{i,m}$ is most convenient if performed at a specified value of control ampere-turns, AT_m. This point should be chosen on the basis of the largest anticipated effective a-c loop width.

Related Magnetic Quantities

As previously stated, there is a correspondence between the core test characteristic curve and an effective major a-c hysteresis loop. In considering core materials, it is convenient to express the core test characteristic values in terms of the normalized magnetic characteristics of this effective hysteresis loop. The relations between the core test measurements and the corresponding magnetic quantities are based on the equations

$$H=\frac{0.4\pi AT}{l_m} \qquad (6)$$

$$\Delta B=\frac{V_i}{2.22fA_c}\times 10^8 \qquad (7)$$

where

l_m = the mean magnetic path length
A_c = effective core cross-section area
ΔB = peak-to-peak flux density change per cycle

The conversion of core test characteristics and test conditions to effective magnetic quantities are as follows.

For core test values

$$B_m-B_r=\frac{V_{i,0}}{2.22fA_c}\times 10^8 \qquad (8)$$

$$H_d=\frac{0.4\pi AT_d}{l_m} \qquad (9)$$

$$\mu_d=\frac{\Delta B_2-\Delta B_1}{\Delta H}=\frac{G_i l_m}{0.888fA_c}\times 10^8 \qquad (10)$$

$$B_m=\frac{V_{i,m}}{4.44fA_c}\times 10^8 \qquad (11)$$

For test conditions

$$H_p=\frac{0.4\pi AT_p}{l_m} \qquad (12)$$

$$\Delta B_1=\frac{V_{i,1}}{2.22fA_c}\times 10^8 \qquad (13)$$

$$\Delta B_2=\frac{V_{i,2}}{2.22fA_c}\times 10^8 \qquad (14)$$

$$H_m=\frac{0.4\pi AT_m}{l_m} \qquad (15)$$

Typical ranges of these normalized magnetic characteristics as measured for a 400-cps test frequency are given in Table I for the common rectangular loop materials used in magnetic amplifiers. The data represented have been accumulated over the past $1^1/_2$ years and are felt to be representative of the 400-cps characteristics at present available in wound-strip toroidal cores of commercial materials 1-mil or 2-mil thick.

Core Shape Effects

It might be expected that B_m, B_r, H_d, and μ_d would be functions strictly of the core material, and independent of core shape. However, it has been found experimentally that μ_d decreases as core build-up or outside-diameter (OD) to inside-diameter (ID) ratio is increased. This effect has been found both from core test and d-c differential permeability measurements. That such an effect should occur appears logical upon reasoning that with a large OD-to-ID ratio there will be a large difference in the magnetizing force applied to the

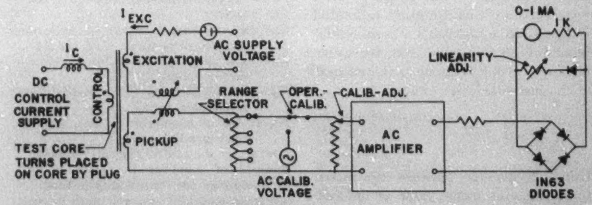


Fig. 2. Block diagram of core tester showing instrumentation for induced voltage measurement

inner and outer portions of the core at a given value of applied ampere-turns. This will tend to increase the differential ampere-turns required to change the entire core from positive to negative saturation. Initial theoretical studies of this effect, on the basis of an assumed idealized hysteresis loop for a thin ring of material, indicate a rather complicated relationship between the effective differential permeability, material magnetic properties, and core OD-to-ID ratio. Experimental determinations of this relationship are complicated by the ever-present variance of magnetic characteristics found between different cores because of material and manufacturing process variables. However, the indications are that for OD-to-ID ratios less than about 1.4 the reduction in μ_d is relatively small for present commercially available materials. Since most of the widely used core sizes fall within this range, the effect will be neglected in the results of this paper.

Core Tester Design

The average sensitive a-c voltmeter is the most difficult portion of the core tester to obtain. The wave form of induced voltage is highly peaked and of a low average magnitude. For small cores it may well be less than 5 millivolts while the ratio of peak to average is between 20 and 50. Average sensitive a-c voltmeters which will measure such a voltage and be truly average sensitive are not commercially available. The best solution found to date has been to use a wide-band high-gain electronic amplifier to raise the level and then to meter with a linearized rectifier-type voltmeter. Care must be taken in the choice of amplifier since phase shift within the frequency spectrum of the induced voltage will introduce an error in the average value measurement. A block diagram of a core tester with such a meter arrangement is shown in Fig. 2.

Placement of the test turns on a core has not been found critical in the case of toroidal cores. Leakage flux is apparently not appreciable, and it is possible to place the necessary turns on the core by means of a simple multipin plug-and-socket arrangement. However, it is necessary to eliminate the air core coupling between the excitation and pickup windings so that the induced voltage will be zero in the absence of a core. This can be accomplished by physically separating the turns on the core, but it is more convenient to compensate for the coupling by the use of an adjustable linear bucking transformer, as shown in Fig. 2.

Another important feature of the core tester design is the loading placed on the core by the control and pickup windings. The effect of this loading will be to reduce the net applied magnetizing force by an amount dependent on the induced voltage v_t and the number of turns which are loaded. Thus the effect of the loading will be to decrease the apparent core gain. An approximate calculation of the gain reduction can be made if it is assumed that the total amount of flux reset during a half-cycle period is proportional to the average control magnetizing force applied during the period. This is only an approximation but will serve to indicate the order of magnitude of loading effects. Considering the effect of the induced voltage on the control current, the average net ampere-turns applied to the core during the reset period will be

$$AT=I_cN_c-\frac{N_c^2\hat{V}_t}{R_c} \tag{16}$$

where I_cN_c is the nominal value of d-c control ampere-turns and R_c is the resistance in series the control winding of N_c turns. The symbol $\wedge$ indicates a half-wave average value, i.e.,

$$\hat{V}=\frac{1}{T/2}\int_0^{T/2}v(t)dt \tag{17}$$

Utilizing the foregoing assumption, the value of AT given by equation 16 will determine the amount of reset, and thus the value of V_t, in a manner predicted by the core test characteristic curve. This relation will be the equation of the linear region of the core test characteristic curve or

$$V_t=G_tAT+C \tag{18}$$

where C=a constant given by

$$C=V_{t,1}-G_1AT_d \tag{19}$$

Substituting the equation 16 value of AT into equation 18 gives

$$V_t=G_tI_cN_c-\frac{G_tN_c^2\hat{V}_t}{R_c}+C \tag{20}$$

The half-wave average and average sensitive a-c meter values of induced voltage V_t are related by $\hat{V}_t=0.9\ V_t$. Solving for V_t

$$V_t=\frac{G_tN_cI_c+C}{1+(0.9G_tN_c^2)/R_c} \tag{21}$$

Thus it is seen that for the loading to have a negligible effect, 0.9 $G_tN_c^2/R_c$ should be very much less than 1. In practice it is found convenient to use inductive rather than resistive isolation of the control source in order to reduce the control power requirements.

Relation of Amplifier Performance to Core Test Data

Practical evidence for the existence of a general correlation between core test results and amplifier operating characteristics is the appreciable increase in yield that resulted when the tester was introduced for the selection of matched groups of toroidal cores for the production of a variety of critical self-saturating magnetic amplifiers. For production use, relatively broad general limits were placed on the first and fourth test points, while the required range and match of values for AT_d and G_t were determined for each individual amplifier design. When this procedure for core selection and matching was introduced on a pilot plant scale, in place of a

Table I. Typical Core Test Magnetic Characteristic Ranges for Various Materials at 400-Cycle Test Frequency

	Test Values				Test Conditions			
Core Material, Toroidal Cores	B_m-B_r, Kilogausses	H_d, Oersteds	μ_d	B_m, Kilogausses	AT_p, Oersteds	ΔB_1, Kilogausses	ΔB_2, Kilogausses	AT_m, Oersteds
Highly oriented 50% nickel-iron, 2-mil strip, Hipernik V	0.9 to 1.4	0.15 to 0.24	15 to 50×10⁴	14.0 to 15.3	2.0	10.5	21	0.359
Nonoriented 50% nickel-iron, 2-mil strip, Hipernik	1.4 to 3.5	0.08 to 0.16	15 to 50×10⁴	12.0 to 14.0	2.0	10.5	21	0.359
Mo-Permalloy and supermalloy-type materials, 1 or 2-mil strip	1.3 to 3.1	0.015 to 0.05	40 to 90×10⁴	5.7 to 8.7	1.0	5.25	10.5	0.135
Oriented 3% silicon-iron, 5-mil strip, Hipersil	1.3 to 3.5	0.39 to 0.54	2.6 to 11×10⁴	14.4 to 17.5	4.0	10.5	21	1.58

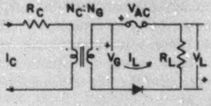


Fig. 3. Basic half-wave self-saturating magnetic amplifier

previous test method based on measurements of the peak differential permeability and coercive force for a major a-c hysteresis loop obtained with sinusoidal magnetizing force, the yield at the first operational test point increased from about 70 to 90 per cent. A continued increase in the yield figure has been observed since the introduction of the tester, partly because of test equipment refinements and better specification of allowable characteristic ranges.

The process of determining the allowable ranges of core characteristics for each new amplifier design during the past $1^1/_2$ years has afforded many opportunities for checking the relation between core test data and amplifier operating characteristics. In the majority of instances at least a qualitative correlation was found, i.e., cores with high values of G_t would result in amplifiers with somewhat higher gains, and also the amplifier bias requirements would fall in order with the core test values of AT_d. However, there were some amplifier designs for which little direct relation between amplifier performance and core test data could be found. In some of these cases it was determined that intrawinding capacitance was introducing an appreciable effect, thus masking core variations. In others there was no doubt that some other facet of core behavior was having an appreciable effect on amplifier performance.

MECHANISM OF AMPLIFIER OPERATION

An analysis of the mechanism of operation of a self-saturating amplifier will serve to illustrate the manner in which core magnetic properties affect amplifier performance and will also indicate the conditions on amplifier design necessary for direct correspondence to the core tester measurements. The basic half-wave circuit in Fig. 3 will be analyzed in detail throughout a full cycle. It is assumed for the sake of simplification that the rectifier is ideal, having zero reverse leakage and a constant forward resistance. Also the control signal source impedance is effectively infinite, the load is purely resistive, and the total load circuit resistance is lumped into R_L. Basic equations expressing amplifier control characteristics in terms of core tester measurements, as well as the conditions necessary for the validity of these equations, will be obtained. These results can then be easily extended to the case of more complicated circuits.

As is shown in Fig. 4(C), at the start of a positive half-cycle of supply voltage the flux will be at the reset level ϕ_0 as determined by the control demagnetizing force applied during the previous half-cycle. As the supply voltage increases in the positive direction, current will flow through the gate winding applying positive ampere-turns to the core. This will increase the net ampere-turns acting on the core, Fig. 4(B), which will produce a flux change and an induced voltage in the gate winding, tending to limit the load current. This action will continue, with the load current being limited to the magnetization current for the core, until the flux reaches a positive saturation. Once the core saturates, the full supply voltage appears across the circuit resistance, or in this case R_L, and the load current will increase to a value many times the magnetizing current.

Toward the end of the positive half-cycle the supply voltage and load current will decrease to a point where the net ampere-turns will not be sufficient to keep the flux at positive saturation. The decrease in flux will induce a voltage in the gate winding, tending to increase the load current. This will increase the load voltage to a value slightly greater than the supply voltage. However, in nearly all magnetic amplifier designs the load impedance is low compared to the unsaturated impedance of the gate winding, and a very small value of induced voltage will be sufficient to prevent further decrease of load current, and thus will keep the flux essentially at positive saturation until the end of the positive half-cycle. This may be seen in Fig. 4(A) where the induced voltage is essentially zero until the supply voltage reverses polarity.

From this it will be seen that the load voltage will be equal to the supply voltage during the positive half-cycle, minus the voltage absorbed by the gate winding as the core flux changes from the reset point ϕ_0 to positive saturation. The average value of gate-winding induced voltage will be related to the flux change. Thus using half-wave average values, as defined in equation 17, the load voltage will, to a close approximation, be given by

$$\hat{V}_L=\hat{V}_{AC}-\hat{V}_G \qquad (22)$$

where

$\hat{V}_L$ = average value of load voltage during the positive half-cycle
$\hat{V}_{AC}$ = average value of supply voltage during the positive half-cycle
$\hat{V}_G$ = average value of gate-winding induced voltage during the positive half-cycle

The corresponding full-wave values of these quantities will be as follows: For the load voltage, the full-wave average d-c value will be one-half of the half-wave average because there will be negligible output during the reset half-cycle. For both the supply voltage and the gate-winding induced voltage, the half-wave average values will be equal in magnitude during each half-cycle and thus the average sensitive a-c meter values will be $\pi/2\sqrt{2}$ or 1.11 times the half-wave average values. Thus equation 22 can be written

$$V_L=0.45(V_{AC}-N_GV_t) \qquad (23)$$

where $N_GV_t=V_G$, since V_t is defined as the average sensitive meter value of the induced voltage per turn.

It is worthy of note that equation 23 is explicitly independent of the magnetizing current that flows as the core is driven to positive saturation. The flow of magnetizing current will change the firing angle and the wave form of the output voltage, but will not affect the average output so long as the core reaches saturation before the end of the positive half-cycle of supply voltage. However,

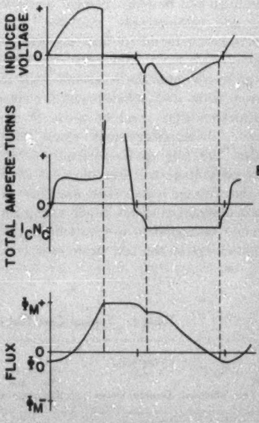


Fig. 4. Typical wave forms for half-wave self-saturating magnetic amplifier

A—Voltage induced in gate winding
B—Total ampere-turns on core
C—Flux in core, negative control ampere-turns applied to core

the induced voltage during the reset of flux can produce a load current similar to magnetizing current which may affect the average output, as will be shown.

At the beginning of the negative half-cycle of supply voltage the flux will be near to positive saturation and there will be a current circulating in the load circuit tending to keep the flux at saturation. As the supply voltage increases in the negative direction the load circuit current will decrease and thus increase the net demagnetizing ampere-turns applied to the core, as is shown in Fig. 4(B). The induced voltage will therefore tend to follow, and in effect be limited by, the supply voltage. (This action has been termed "backfiring" in the sense that current flows in the load circuit at a time when the supply voltage is of a polarity to put a back voltage on the rectifier.) The induced voltage will continue to be limited by the supply voltage until a point is reached where the maximum net demagnetizing force available, (corresponding to the applied negative control ampere-turns) will not cause a flux rate of change great enough to induce a gate-winding voltage equal to the supply voltage. At this time the load current will tend to reverse, the rectifier will block, and the flux will continue to reset at a natural rate determined by the influence of the control demagnetizing force on the core.

For cores with a relatively short time lag for a change of magnetization, the flux will have reset to a point on the d-c hysteresis loop before the end of the half-cycle. However in the case of the metallic cores commonly used in magnetic amplifiers there will be a range of control for which the flux will still be changing at the end of the negative half-cycle. For this latter case backfiring will again occur when the induced voltage equals the supply voltage. This effect is seen in Fig. 4 just before the end of the negative half-cycle.

A comparison of the amount of flux reset obtained in amplifier and core tester operation of a core can now be made. It is seen that for the case of cores with a negligible time lag for a change of magnetization, the amount of flux reset will be exactly the same in core tester or amplifier operation with a given value of control ampere turns. For cores with a long magnetization time lag, however, there will be a difference in the amount of flux resetting dependent on the portion of the negative half-cycle during which backfiring occurs. This is a function of both amplifier design and core characteristics.

AMPLIFIER EQUATIONS

For the case of amplifier designs for which the amount of flux reset for a given value of control current will be the same as in the core tester, the total flux change per cycle and thus the value of induced voltage per turn V_t will be the same in each. Equation 18, relating V_t and the control ampere-turns for the core tester, can be substituted into equation 23 for the half-wave amplifier giving for the transfer curve of a half-wave amplifier

$$V_L=0.45(V_{AC}-G_tN_GAT-C) \qquad (24)$$

The gain K of this amplifier will be dV_L/dAT or

$$K=-0.45G_tN_G \qquad (25)$$

Extension of these results to the various full-wave circuit configurations can be easily accomplished if each core in the full-wave circuit can be assumed to operate in the same manner as in a half-wave circuit. Lord[2] has shown this to be the case for the bridge and center-tap circuit but not for the doubler circuit in the case of typical circuit designs and core characteristics. Actually the action of the load circuit in limiting the induced voltage during reset will be somewhat different in each of the circuits and the flux reset may or may not be the same as in core tester operation, dependent on the core characteristics and amplifier design. For those full-wave circuits with high control circuit impedance, for which the requirements are met, the output each half-cycle will be as given by equation 22 and thus for the case of d-c output

$$V_L=0.9(V_{AC}-G_tN_GAT-C) \qquad (26)$$

and for a-c output, as from a doubler circuit

$$V_L=V_{AC}-G_tN_GAT-C \qquad (27)$$

The effects of rectifier leakage can be computed for each type of circuit if it is assumed that the amount of flux reset is proportional to the average net ampere-turns applied to the core during the reset period. It is to be shown that this is not an accurate assumption and therefore such a calculation is useful only for estimating the order of magnitude of the rectifier leakage effect.

EXPERIMENTAL EXAMPLE

The relation between core test measurements and amplifier performance and, in particular, the effect of amplifier design on the amount of backfiring will be further illustrated by an experimental example. A half-wave self-saturating amplifier with essentially infinite-control circuit impedance, such as is shown in Fig. 3 was constructed with the following core and design features.

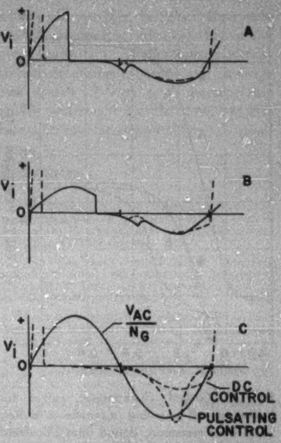


Fig. 5. Induced voltage wave forms (for core of Fig. 6 with —1.3 control ampere-turns applied). Solid lines, core operating in amplifier. Dashed lines, core operating in core tester. Operating conditions:

A—D-c control current, $V_{AC}=75$ volts
B—D-c control current, $V_{AC}=40$ volts
C—Pulsating control current, reference supply voltage shown

Characteristics of Example Core

Size—3/4-inch ID by 1-inch OD by 1/4-inch height
Material—oriented 50 per cent iron-nickel, Hipernik V
Construction—toroidal core, wound from 2-mil-thick strip with 84 per cent space factor
Core test values at 400 cps:
  $V_{t,0}$: 0.75 by 10^{-8} volts per turn
  AT_a: -1.30 ampere-turns
  G_t: -0.088 volts per turn per ampere-turn
  $V_{t,m}$: 43.5 by 10^{-8} volts-turn

Example Amplifier Design

Gate winding—1,500 turns
Self-saturating rectifier—two type-*6AL5* vacuum-tube rectifiers in parallel (four plates)
Total effective load resistance, R_L—1,100 ohms
Nominal saturation voltage at 400 cps—60 volts rms

Control transfer curves for the example amplifier under various operating conditions are shown in Fig. 5. Also shown, in dashed lines, are the transfer curves as calculated from the core test charac-

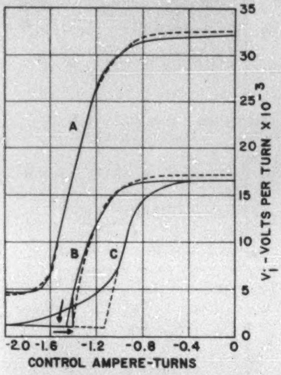


Fig. 6. Control characteristic curves for half-wave self-saturating magnetic amplifier. Solid lines measured, dotted lines calculated from core test data. Operating conditions:

A—D-c control current, $V_{AC}=75$ volts
B—D-c control current, $V_{AC}=40$ volts
C—Pulsating control current, $V_{AC}=40$ volts

teristic curve of Fig. 7 with the use of equation 24.

It will be noted from curves *A* and *B* of Fig. 6 that there is almost a perfect correspondence between measured and calculated values of amplifier output for operation with a supply voltage of 75 volts, while the correspondence is not as good with a supply voltage of 40 volts. The reason for this is evident from the induced voltage wave forms shown in Fig. 5. When the amplifier is operated at 75 volts supply voltage, the induced voltage in the gate winding during the reset period does not tend to exceed the supply voltage for a large portion of the period, and the amount of flux reset obtained will be approximately the same in the core tester or amplifier. However, when the amplifier is operated at 40 volts' supply voltage, the voltage induced in the gate winding during reset is effectively limited to the value of supply voltage during an appreciable portion of the period, and there is a difference in the amount of flux reset obtained under amplifier and core tester operation.

PULSATING CONTROL CURRENT

Practical applications often require that a magnetic amplifier be controlled by a pulsating, rather than a pure d-c control current. As the experimental example demonstrates, the effect of pulsating control currents on an amplifier with relatively high control circuit impedance can be twofold. First, there is a difference in the amount of flux reset obtained in a core when a constant demagnetizing force is applied for one half-cycle of the test frequency or when a sinusoidal pulse of demagnetizing force, having the same average value, is applied for the same period of time. Second, the voltage induced in the gate winding of an amplifier during the flux reset period will tend to have a higher peak value when a pulsating control demagnetizing force is applied, and this can cause a drastic amount of backfiring and lack of correspondence between core tester results and amplifier performance.

These two effects of the use of pulsating control current are illustrated in Figs. 5, 6, and 7 for the case of control current obtained from full-wave rectified supply voltage. The effect on the core test characteristic curve is shown in Fig. 7. The major effect of the pulsating control is to produce a lateral shift of the curve. The slope is changed but little; from $G_t=-0.088\ v/T$ per A-T for pure d-c control to $G_t=-0.095\ v/T$ per A-T for full-wave rectified control current. Comparative tests on other cores indicate that the gain change may be either slightly positive or negative but that it is usually minor as compared to the shift in effective loop width or bias requirement.

The cause for the second effect, that of increased backfiring, is apparent from the wave form of induced voltage, Fig. 5(C). It is apparent from this wave form that amplifier performance should correlate very well with core test measurements, made under control current wave form conditions, until a point is reached where the peak of induced voltage produces backfiring. This is demonstrated in Fig. 6, curve *C*, where the measured curve is seen to break away at a point part way down the transfer curve. An investigation of the circuit operation at this point showed that the peak of induced voltage in the gate winding was just equal to the supply voltage.

The implications of this control current wave form sensitivity are of importance. First, the anomalous results sometimes encountered in attempting to control a high control circuit impedance amplifier with the highly pulsed output of a previous stage, are explained in terms of the presence of excessive backfiring during the period of the pulse. Second, in the case of full-wave amplifiers with relatively low control circuit impedance, the voltage induced in the control winding during reset is forced to be equal to the voltage induced in the control winding of the opposite core as it is driven to saturation by its gate circuit.[6] Under these conditions there will be no appreciable backfiring but the net control current will have a pulsating wave form dependent on the magnetization current of the core being reset.[6] Obviously, the correlation with core tester data obtained with pure d-c control current will be poor under these conditions. However, if the results obtained with the use of full-wave rectified control current can be extrapolated, it appears the amplifier gain in the low control circuit impedance case should be related to the core test gain measure. Experimentally, it has been found that a qualitative relation does exist.

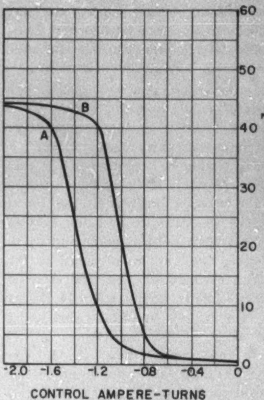


Fig. 7. Core test characteristic curves (for core of Figs. 5 and 6). Operating conditions:

A—D-c control current
B—Pulsating control current

Conclusions

For simple self-saturating magnetic amplifiers, constructed with cores having negligible eddy-current type of effects, such as the rectangular loop ferrite cores, the control characteristic curve will be an accurate reflection of the left flank of the hysteresis loop for the cores used in the amplifier. That is, during the reset half-cycle of supply voltage with respect to a particular core, that core will be reset to a point on the d-c hysteresis loop corresponding to the control demagnetizing force applied. On succeeding positive or gating half-cycles the integral of the voltage induced in the gate winding, corresponding to the flux change back to saturation, would then be effective in reducing the load voltage.

However, in the case of metallic cores, and even for very thin thicknesses of material, it is found that the time lag for a change of magnetization, as produced by eddy currents or other similar phenomena, plays a major role in determining the operation of a self-saturating amplifier. The period of time available for flux reset, equal to a half-cycle of the operating frequency, is not sufficient to obtain complete flux reset back to a point on the d-c hysteresis loop. The core test method which has been described takes the time lag of magnetization into account by measuring the amount of flux reset obtained when a constant demagnetizing force is applied for a period of one half-cycle of the operating frequency.

A further investigation of amplifier operation discloses that the core tester gives a realistic approximation to amplifier operation for some designs but that for others another factor is affecting amplifier operation. This is a limiting of the flux rate of change brought about by a limiting of the induced voltage which can appear in the gate winding during the reset period. This limitation of flux rate of change, coupled with the time lag of magnetization, results in less reset than predicted by the core tester.

Finally, it is found that the wave form of the control demagnetizing force has a two fold effect on amplifier operation. In the case of flux reset at a natural rate, the wave form of demagnetizing force will directly influence the total amount of reset obtained during the half-cycle period. But, in addition, the application of pulses of demagnetizing force to a core operating in an amplifier may not be fully effective because of the limitation of the flux rate of change by the gate circuit.

It is evident that a modification of the core tester to introduce the limiting action on the flux rate of change and the control current wave form corresponding to a particular amplifier design would make possible the accurate prediction of core performance in that amplifier design. However, this would require a different tester design for each type of amplifier design. Such a procedure might be justified in the case of large-volume production of a highly critical amplifier. But the advantages of a general-purpose magnetic amplifier core test method would be lost.

The present core test method, though not perfectly accurate in predicting the operating characteristics of some types of magnetic amplifier design, has been found accurate in selection of cores which will have the characteristics required for the proper operation of nearly all amplifier designs. Extensive production use has indicated that in the great majority of applications it is possible to determine experimentally a range and match of core test characteristic values which specify the core requirements for satisfactory amplifier operation.

References

1. Predetermination of Control Characteristics of Half-Wave Self-Saturated Magnetic Amplifiers, **Henry Lehmann.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 2097–2103.

2. The Influence of Magnetic Amplifier Circuitry Upon the Operating Hysteresis Loops, **H. W. Lord.** *AIEE Transactions*, vol. 72, pt. I, July 1953, pp. 721–28.

3. Flux Resetting Characteristics of Several Magnetic Materials, **Hoobert Huhta.** *AIEE Transactions*, vol. 73, pt. I, May 1954, pp. 111–14.

4. Magnetic Amplifier Gapless Core Tests, **J. R. Conrath.** *Electronics*, New York, N. Y., vol. 25, Nov. 1952, pp. 119–121.

5. Ferromagnetism (book), **R. M. Bozorth.** D. Van Nostrand and Company, Inc., New York, N. Y., 1951.

6. Theory of Magnetic Amplifiers with Square-Loop Core Materials, **H. F. Storm.** *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 629–40.

7. Dynamic Hysteresis Loop Measuring Equipment, **H. W. Lord.** *AIEE Transactions*, vol. 72, pt. I, Sept. 1952, pp. 269–72.

8. Dynamic Hysteresis Loops of Several Core Materials Employed in Magnetic Amplifiers, **H. W. Lord.** *AIEE Transactions*, vol. 72, pt. I, Mar. 1953, pp. 85–88.

9. Pulse Response Characteristics of Rectangular-Hysteresis Loop Ferromagnetic Materials, **Joseph Wylen.** *Ibid.*, pp. 648–56.

10. Evaluation of Core Materials for Magnetic Amplifier Applications, **R. D. Teasdale, H. R. Brownell.** 1954 National Convention Record, pt. III, March 1954, pp. 56–65.

◆

Discussion

Clement L. Boyajian (General Electric Company, Schenectady, N. Y.): This is a timely paper on the measurement of the magnetic characteristics of cores. Of major interest is the correlation between the characteristics measured on the core tester and the actual amplifier characteristics.

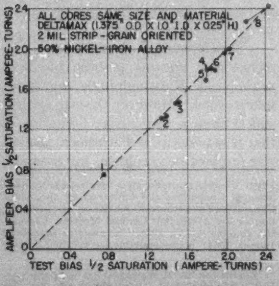


Fig. 8. Amplifier bias for 1/2-saturation voltage versus tester bias for 1/2-saturation voltage

Considerable work has been done in our laboratories on obtaining data to determine this correlation, and a presentation of a portion of the findings may be enlightening at this time.

The core tester used was a constant current reset tester similar to the one referred to in Fig. 8. The magnetic-amplifier circuit used was a full-wave, center-tapped circuit employing electronic diodes. The correlation resulting is shown in Figs. 8 and 9. The two plots are the most pertinent of the data compiled.

The data were obtained by choosing eight matched pairs of cores, all of the same size and material. They were of Deltamax (*4635-D2*) 2-mil strip grain-oriented 50-per-cent nickel-iron alloy, having rectangular hysteresis loops (size 1.375-inch OD by 1-inch ID by 25-inch height). In Figs. 8 and 9, pairs are numbered 1 to 8, according to width of loop from narrowest to widest, respectively.

The pairs were chosen so that their widths of loop were as diversified as possible. They represent a realistic cross section of cores which were received from a given manufacturer over a period of months. It is important to note that for this material a wide variation in coercive force was encountered. The measured coercive forces varied approximately 2.5 to 1.

These cores were first tested in the core tester measuring the bias necessary to reset the core to point *A* and the induced voltages corresponding to resets to points B_1 and B_2 as shown on the *B-H* loop in Fig. 10. Point *A* was determined by measuring the maximum or saturation value of induced voltage for a specific maximum bias. Then the bias current necessary to induce one-half of this saturation voltage was measured. This is re-

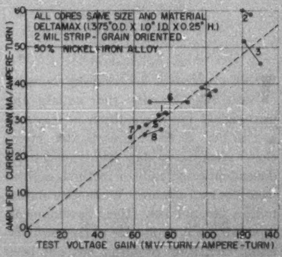


Fig. 9. Amplifier current gain versus tester voltage gain

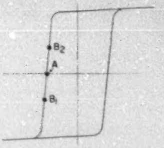


Fig. 10. Illustrative B-H loop

ferred to as the test bias for 1/2-saturation voltage. This bias point is representative of the coercive force of the core, although not a measure of it. The maximum bias current and the excitation current were the same for all cores.

To obtain the core voltage gain on the core tester, the induced voltage was measured for a specific plus-and-minus bias current, straddling the bias of point A. This represents resets to points B_1 and B_2.

The cores were then wound and their amplifier characteristics measured for high-impedance control operation. A ratio of control turns squared to control resistance (N^2/R) equal to 10 was considered high-impedance control. For the amplifier measurements, the supply voltage was kept constant for all cores at a value of approximately 75 per cent of the average saturation voltage of the cores.

The amplifier bias current necessary to induce in each core a voltage exactly equal to the 1/2-saturation voltage measured on the tester was measured. This is referred to as the amplifier bias for 1/2-saturation voltage. Except for the slight discrepancy caused by the greater peak excitation current of the amplifier, this value of amplifier bias current would cause the core to reset to the same point on the B-H loop as it had done on the core tester. For perfect correlation the bias ampere-turns measured in each case should be equal. The plot of Fig. 8 shows this to be very nearly true, resulting in excellent correlation for the high-impedance control condition. However, as would be expected, the correlation was not as good for lower impedance control conditions.

To obtain the amplifier gain data, the bias currents necessary to induce exactly the same voltages as measured on the tester for points B_1 and B_2 were measured, as well as the corresponding amplifier load currents. This insured gain measurement at the same points of operation for both cases.

The correlation of gain measurements was fairly good, as can be seen in Fig. 9. For perfect correlation, all points would fall on a straight line passing through zero.

R. W. Roberts: I wish to thank Mr. Boyajian for his findings on the correlation between core tester measurements and actual amplifier performance. His results are quite similar to those we have obtained from like experiments. The test points which he uses to specify the important features of the core test characteristic curve are an interesting alternative to those I used. Certainly the value of control ampere-turns necessary to induce one-half of saturation voltage is of basic interest, since this value is, in a certain sense, analogous to a coercive force. However, from a purely practical standpoint of testing cores for use in magnetic amplifiers it would seem that a measurement of the control ampere-turns necessary to induce some fixed value of voltage would be more directly useful and would be an easier value to obtain.

A few additional remarks on the known causes of discrepancies between core tester and amplifier characteristics will perhaps be of value to others who wish to investigate the core test methods. It has been found generally true that, as demonstrated by Boyajian's results, the best correlation exists for effective loop width or bias measurements, while the gain figures are in poorer agreement. There are several apparent causes for this. First, gain measurements in both core tester and single-ended amplifiers involve measurements of a small change in a relatively large value of control or bias current. Thus, normal random metering errors will produce a greater percentage variance in gain measurements than in bias measurements. Second, mechanical strain or core damage will affect core gain much more than the bias value. In a great many instances apparently poor correlation has been found to be caused by slight core damage because of rough handling during the process of placing windings on the core for testing in an amplifier.

When large numbers of turns are required for amplifier operation, the distributed intrawinding capacitance can have an appreciable effect on amplifier performance. This effect is usually apparent as a change in both amplifier gain and bias requirement, and can, of course, lead to very poor correlation with core tester predictions. Often it is quite difficult to know when the distributed winding capacitances are having an appreciable effect on amplifier characteristics, and the variations in the distributed capacitance from core to core can completely mask any correlation with core tester measurements.

Finally, when these sources of discrepancy are eliminated, it is still found that an occasional core will show a poor correlation. It is possible to find two cores at a given test frequency, with almost identical core test characteristics, yet which exhibit quite different characteristics when operated in an amplifier. The cause has been found to lie in the different shapes of the wave forms of induced voltage during the reset period. Since the core tester is sensitive to the integral of induced voltage during the reset period, it cannot detect slight variations in the wave form of this voltage. An added bit of information necessary to detect wave form variations for the purposes of accurate core grading and matching can be obtained by either observing the reset induced voltage wave form with an oscilloscope, or by making tests at two or more frequencies.

A Symbolic Method for Synthesis of 2-Terminal Switching Circuits

D. ZEHEB
ASSOCIATE MEMBER AIEE

W. P. CAYWOOD
ASSOCIATE MEMBER AIEE

THE problem of synthesizing minimal switching circuits that satisfy any preassigned input condition is one that has of late received considerable attention. Analysis to date has tended to be dependent principally on methods derived from Boolean algebra. While such mathematics are not new, their development, especially as is applicable to the combinational switching problem, has not been comparable to the more familiar algebra of real numbers.

The engineer faced with the problem of designing a specific switching circuit has a set of conditions which the circuit is desired to fulfill. These conditions are easily translated from their verbal form into Boolean algebraic expressions. The next step is to simplify these expressions until they clearly indicate the most economical circuit for each specific case. Although the basic rules for the manipulation of Boolean algebraic expressions are quite simple, the method of arriving at the most economical circuit is not necessarily obvious. The difficulty seems to be that the designer has no immediate way of knowing which theorem of those available to use and knowing to which of the expressions to apply it next.

A step forward in the solution of this increasingly pressing problem has been made by Dr. Karnaugh in developing his map method.[1] However, the map method suffers from a very severe limitation in that it requires an n-dimensional parameter (an area or a volume, for example)

Paper 54-394, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted June 8, 1954; made available for printing August 19, 1954.

D. Zeheb and W. P. Caywood are with the Carnegie Institute of Technology, Pittsburgh, Pa.

The research described was supported in part by the Office of Naval Research under Contract *N7-onr-30306*.

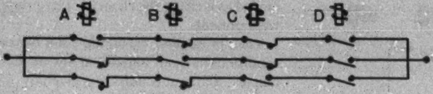


Fig. 1. Representing a circuit as three parallel paths

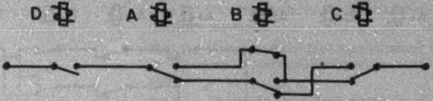


Fig. 2. Simplified form of circuit represented in Fig. 1

for representation of all of the possible combinations. Practicability limits the representation to six elements on a 3-dimensional map.

In contrast, in the presently described method the combinations are represented by "chains" of symbols, resulting in no limitations on the complexity of the case treated. Further, the steps in the present method are explicitly defined in advance within certain choices, and a pursuance of these steps will lead to a final circuit, including the use of transfer contacts when desirable.

The Symbolic Method

DESCRIPTION AND SYMBOLS

This method uses three types of contacts: make, break, and transfer contacts. A make contact is denoted by a letter with a connection on each side, $-A-$, a break contact by a primed letter which has a connection on each side, $-A'-$, and a transfer by a letter with a dash on top, one connection on one side and two connections on the other, $-\bar{A}<$ or $>\bar{A}-$. The upper connection is the make contact and the lower one is the break contact.

In all Boolean algebraic expressions in this paper, multiplication stands for the logician's "and" (intersection) and addition for the logician's "or" (union). An arrangement of the elements which provides a closed path between input and output is called a term. The sum of such terms of a given problem is called an input condition.

For example, if it is required that a path be closed between input and output when relay A is operated and B and C are released, or when B and C are operated and A released, the input condition is given by two terms

$$f(x)=AB'C'+A'BC$$

It is assumed here that a make contact closes when its relay is energized and a break when the relay is not energized.

PROCEDURE

The method may best be illustrated by means of examples. Suppose an input condition is given by

$$f(x)=AB'C'D+A'BC'D+A'B'CD$$

The conventional methods of designing the corresponding circuit are two: inspection and formal methods.[2]

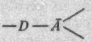
1. Designing by inspection means in this case representing the circuit as three parallel paths, Fig. 1, and searching by inspection for contacts that can be combined. D is obviously such a contact.

2. In the case of designing by formal methods, Boolean algebraic rules are applied to $f(x)$ to get a simplified form. As is described in the preceding section, there remains the problem of which rule to apply where.

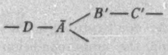
In the symbolic method the design procedure is started by looking for elements common to all terms. D is such an element, so D is written in its symbolic form

$$-D-$$

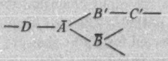
Next, since there are no more common elements, any one of the rest can be chosen, e.g. A. Since A appears as both a make and a break, it can be written as a transfer contact



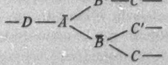
There is but one term with A (unprimed) in the example. This term is written at the make connection of A



Of the remaining terms, with A', one has a B and the other a B' so a transfer is again used for B

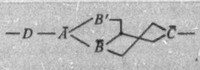


The remainder of the terms are completed and the following is obtained.



This stage of the symbolic diagram will be called a developed tree stage.

Since a 2-terminal network is being described, the end connections of the tree all go to a common point, namely the output terminal. So the following step is to close the branches of the tree, working from the output terminal backwards. Noticing that C appears both as a make and a break, a transfer is used for C and the make of $\bar{C}$ is connected to the break and $\bar{B}$ and the break of $\bar{C}$ to the make of $\bar{B}$ and to B'



In this simple example it is not necessary to go any further. It is now a matter of inspection to design the corresponding circuit; Fig. 2. By starting the design procedure from a different relay a similar circuit could be arrived at with an interchange of relays, but with the same number of contacts; see Fig. 3.

As a further illustration of the method, a problem presented in reference 2 will be solved. The statement: "Eight relays are divided into two equal groups A and B of four relays each. The relays of group A are designated (A_1) (A_2) (A_3) (A_4) while those of group B are designated (B_1) (B_2) (B_3) (B_4). Design a contact network to light a lamp when the designation number of an operated relay in group A is greater than the designation number of an operated relay in group B. The lamp should light only when one and only one relay in each group is operated."

THE SOLUTION

First the verbal problem is translated into a Boolean algebraic form

$$\begin{aligned}f(x)=\ &A_1'A_2A_3'A_4'B_1B_2'B_3'B_4'+\\&A_1'A_2'A_3A_4'B_1B_2'B_3'B_4'+\\&A_1'A_2'A_3A_4'B_1'B_2B_3'B_4'+\\&A_1'A_2'A_3'A_4B_1B_2'B_3'B_4'+\\&A_1'A_2'A_3'A_4B_1'B_2B_3'B_4'+\\&A_1'A_2'A_3'A_4B_1'B_2'B_3B_4'\end{aligned}$$

In looking for common elements it can be seen that A_1' and B_4' are such.

The developed tree stage is

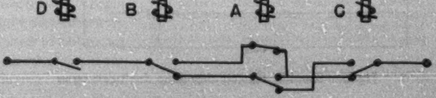


Fig. 3. Alternative simplified form of circuit represented in Fig. 1

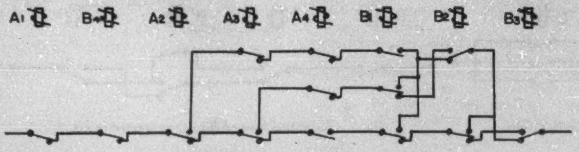
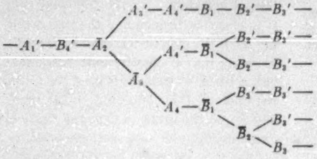
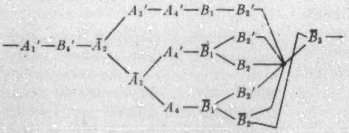
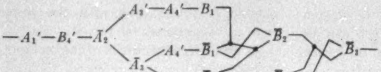

A1
B4
A2
A3
A4
B1
B2
B3

Fig. 4. Circuit to light a lamp when designation number of operated relay in group A is greater than the designation number of operated relay in group B


$A_3'—A_4'—B_1—B_2'—B_3'—$
$—A_1'—B_4'—\bar{A}_2$
$A_4'—\bar{B}_1$
$B_2'—B_3'—$
$B_2—B_3'—$
$\bar{A}_3$
$B_2'—B_3'—$
$A_4—\bar{B}_1$
$B_3'—$
$\bar{B}_2$
$B_3—$

Closing the branches backwards


$A_3'—A_4'—B_1—B_2'$
$—A_1'—B_4'—\bar{A}_2$
B_2'
$\bar{B}_3—$
$A_4'—\bar{B}_1$
B_2
$\bar{A}_3$
B_2'
$A_4—\bar{B}_1$
$\bar{B}_2$


$A_3'—A_4'—B_1$
$—A_1'—B_4'—\bar{A}_2$
$\bar{B}_2$
$A_4'—\bar{B}_1$
$\bar{B}_3—$
$\bar{A}_3$
$A_4—\bar{B}_1$
$\bar{B}_2$

The corresponding circuit is given by Fig. 4.

The "Don't-Care" Terms

So far cases have been considered where the input condition is rigid. That is, a path between input and output should be closed for given conditions and should not be closed for any other possible combination of the variables. In practice, however, somewhat more relaxed requirements are often encountered, i. e., a path should be closed for some terms and should not be closed for others, and may or may not be closed for the remaining combinations. The terms for which the path may or may not be closed are called don't-care terms.

In cases where don't-care terms exist there arises a problem as to which of those terms to assume as corresponding to a closed path and which to an open path, so as to get the most simplified circuit. The symbolic method described in this paper is capable of advantageously using such don't-care terms. The procedure is as follows:

First the don't-care terms are disregarded and the input condition is developed as before until it is in its final stage. Next the don't-care terms are considered and they are developed as if they were the input condition (to result in closed paths). The two diagrams are then inspected for paths that are similar in all elements but one. Such an element can be eliminated in the considered path. A path in the don't-care diagram can be used as many times as is advantageous. Also, an element absent in any of the don't care paths can be assumed to be included as if either the element or its prime were present.

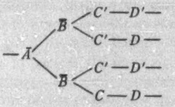
The procedure including don't-care conditions is now illustrated by an example. An input condition is given by

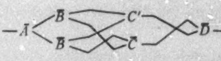
$$f(x)=A'BC'D'+ABC'D'+AB'C'D+A'B'CD$$

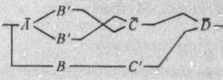
The following don't-care terms also apply

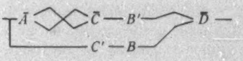
$$AB'CD'+A'BCD'+ABCD'+A'BCD+ABCD+AB'CD$$

As previously explained, $f(x)$ is first developed


$\bar{B}$
$C'—D'—$
$C'—D—$
$—\bar{A}$
$\bar{B}$
$C'—D'—$
$C—D—$

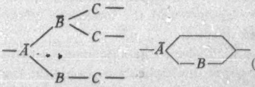

$—\bar{A}$
$\bar{B}$
C'
$\bar{B}$
$\bar{C}$
$\bar{D}—$


$\bar{A}$
B'
B'
$\bar{C}$
$\bar{D}—$
B
C'


$\bar{A}$
$\bar{C}—B'$
$\bar{D}—$
$C'—B$

(1)

The two B contacts are slid until they reach a common point. This is permissible, as can be shown by the commutative law of Boolean algebra. The don't-care terms are now considered.


$\bar{B}$
$C—$
$C—$
$—\bar{A}$
$B—C—$
$—\bar{A}$
B

(2)

Equations 1 and 2 will now be inspected for paths that are similar except for one element. For clarity this is done step by step as follows:

In equation 1 there is a path

terminal—C'—B—D'—terminal

and in equation 2

terminal—A—(B)—C—(D')—terminal

and

terminal—A'—B—C—(D')—terminal

which is equivalent to a path

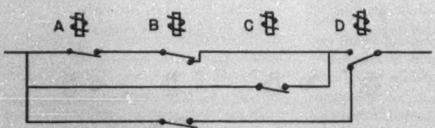


Fig. 5. Simplified circuit after using don't-care terms

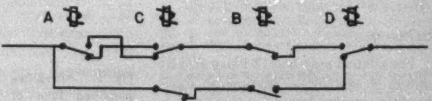
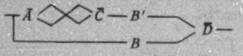


Fig. 6. Circuit of Fig. 5 without use of don't-care terms

terminal—B—C—D'—terminal

so C in this path can be eliminated.

There now results



(3)

In equation 3 is a path

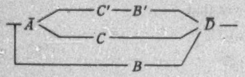
terminal—A'—B'—C—D—terminal

and in equation 2

terminal—A'—B—C—(D)—terminal

so B in this path can be eliminated.

Then



(4)

A path in equation 4 is

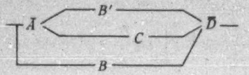
terminal—A—C'—B'—D—terminal

and in equation 2

terminal—A—C—(B')—(D)—terminal

C in this path can be eliminated.

Now



(5)

A path in equation 5 is

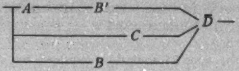
terminal—A'—C—D—terminal

and in equation 2

terminal—A—C—(D)—terminal

A can be eliminated in this path.

Then



The simplified circuit is given in Fig. 5. It is of interest to compare this circuit to that obtained from equation 1 (i.e., without using the don't-care terms); this is shown in Fig. 6. There is a saving of four contact springs.

Conclusion

The steps involved in designing a circuit according to the symbolic method are simple and straightforward. They are here summarized:

Step 1. Translation of the problem from its verbal form into Boolean algebraic expressions.

Step 2. Looking for elements common to all terms. (This step is not a necessity since these elements will combine later anyway.)

Step 3. Developing a tree starting from one of the terminals.

Step 4. Closing the tree backwards from the other terminal.

Step 5. Developing the don't-care terms in the same manner.

Step 6. Comparing the diagrams obtained in steps 4 and 5 and simplifying as much as possible.

Step 7. Designing the actual circuit from the last obtained diagram. Steps 5 and 6 apply only in the case where don't-care terms exist.

The method described in this paper is not claimed to lead always and at first try to the circuit with the least number of elements, nor is it claimed that the method leads to a unique solution. However, as has been pointed out elsewhere,[1] the most economical circuit is not always the one with the least number of elements. In fact the criterion for "most economical circuit" involves many factors besides the number of elements. It is therefore important for the designer to possess a method which will enable him to find a number of equivalent circuits which are near minimal from which to pick the best one. It is hoped that the method described in this paper is of help in this respect.

References

1. THE MAP METHOD FOR SYNTHESIS OF COMBINATIONAL LOGIC CIRCUITS, **M. Karnaugh.** *AIEE Transactions*, vol. 72, pt. I, Nov. 1953, pp. 593–99.

2. THE DESIGN OF SWITCHING CIRCUITS (book), **William Keister, A. E. Ritchie, S. H. Washburn.** D. Van Nostrand Company, New York, N. Y., 1951, chap. 6.5.

Harmonic Analysis for Nonlinear Characteristics

LAUREL J. LEWIS
MEMBER AIEE

Synopsis: Harmonic analysis may be carried out for nonlinear devices by the use of a set of polynomial functions applied directly to the nonlinear characteristic instead of by the more conventional use of trigonometric functions. Higher accuracy may sometimes be achieved in point-by-point calculations. A special method of representation shows how the form of the nonlinear characteristic affects its response. A method is presented for using measurements of amplitudes to obtain analytical expressions for the nonlinear function.

THE application of a nonlinear device such as a vacuum tube, an iron-cored inductance, or a nonlinear resistor frequently requires that a harmonic analysis be made to determine the frequency components in the output for a periodic driving function. While the basis of such an analysis in terms of a Fourier series is well understood,[1] such an analysis is usually expressed in the time domain. It may not be immediately apparent how the form of the nonlinear characteristic affects the amplitude of the frequency components. It is the purpose of this discussion to present a method for making a frequency analysis directly from the nonlinear characteristic. It will then be possible to consider how the shape of the characteristic affects the response, and also how measurements of the response can be used to obtain an analytical expression for the nonlinear characteristic.

Representation of Characteristics

Fig. 1 shows a relatively simple case typical of a number of applications. The nonlinear characteristic $f(x)$ might represent, for example, the plate current of a vacuum tube in which the grid signal x is varying sinusoidally with time. The current variation is a function of time (hereafter expressed as an angular variable θ) which is periodic but nonsinusoidal.

A nonlinear function may be represented in a variety of ways, depending on the requirements. A familiar form is the Taylor series

$$f(x)=a_0+a_1x+a_2x^2+a_3x^3+\ldots+a_nx^n+\ldots \qquad (1)$$

where the coefficients are related to the value of the function and its derivatives

---

Paper 54-512, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted March 22, 1954; made available for printing August 24, 1954.

LAUREL J. LEWIS is with the University of Washington, Seattle, Wash.

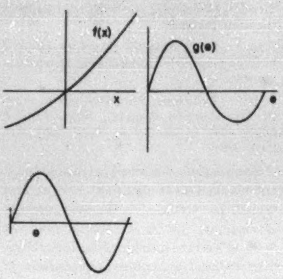


Fig. 1. Example of input and response for nonlinear characteristic

at a point. For the series in equation 1

$$a_n = \frac{1}{n!}\left[\frac{d^n}{dx^n} f(x)\right]_{x=0}$$

and the expression may be considered as an "expansion about a point." The form of representation is restricted to analytic functions, and may in certain cases involve an infinite number of terms.

Another means of representation is in terms of a polynomial. Thus

$$f(x) = c_0 + c_1 x + c_2 x^2 + \ldots + c_n x^n \qquad (2)$$

where n is now finite. This could be a Taylor series if the representation is exact; as an approximation to a physical phenomenon, one might use the first n terms of the Taylor series, that is, by using $c_n = a_n$. More probably, in utilizing a polynomial as an approximation, the coefficients c_n would be adjusted to provide the "best" fit, in accordance with an appropriate criterion. This is a process of curve fitting. One possible criterion is that the coefficients be selected so as to establish coincidence between the actual characteristic and the approximating function at a finite number of selected points. The representation would then be an "expansion in a finite number of points."

The harmonic analysis of the output for a nonlinear system requires a representation by a Fourier series. Thus, for the illustrative case in Fig. 1, let

$f(x)$ = the function whose representation is to be determined
$x = \sin\theta$
= the specific variation required by the analysis (although other variations would be possible)
$g(\theta) = f(\sin\theta)$
= the function whose frequency components are required

Using the trigonometric form of the Fourier series:

$$g(\theta) = \frac{A_0}{2} + \Sigma[A_n \cos(n\theta) + B_n \sin(n\theta)] \qquad (3)$$

where

$$A_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(\theta)\cos(n\theta)d\theta$$

$$B_n = \frac{1}{\pi}\int_{-\pi}^{\pi} g(\theta)\sin(n\theta)d\theta$$

The Fourier expansion constitutes an expansion over an interval, and utilizes information on all parts of the characteristic; it is not restricted to analytic functions.

The relation between the coefficients A_n and B_n of $g(\theta)$ and the coefficients of $f(x)$ depend on the nature of $f(x)$ and also on the way the coefficients of $f(x)$ are to be specified.

Determining the Fundamental Coefficient in the X Domain

As indicated by equation, 3 the Fourier coefficients may be evaluated in the θ domain. Partly to illustrate the general relationship and also to utilize the characteristic of $f(x)$ without having to specify its form, the coefficient for the fundamental frequency term will be expressed in terms of x.

The following substitutions will be evident from Fig. 2. Starting from equation 3

$$B_1 = \frac{1}{\pi}\int_{-\pi}^{\pi} g(\theta)\sin\theta d\theta$$

The integration may be replaced by

$$B_1 = \lim_{\Delta\theta\to 0}\frac{1}{\pi}\sum g(\theta_i)\sin\theta_i\Delta\theta_i$$

From the given variation

$$x = \sin\theta \qquad (4)$$

it follows that

$$\sin\theta_i = x_i$$

and that

$$g(\theta_i) = f(\sin\theta_i) = f(x_i)$$

It is evident from equation 4 that uniform increments $\Delta\theta_i$ do not correspond to uniform increments Δx_i. The relation between increments (neglecting higher-order terms) is

$$\Delta\theta_1 = \left.\frac{d\theta}{dx}\right]_{x_i}\Delta x_i$$

$$= \left.\frac{d}{dx}(\sin^{-1}x)\right]_{x_i}\Delta x_i$$

$$= \frac{1}{\sqrt{1-x_i^2}}\Delta x_i$$

Therefore

$$B_1 = \lim_{\Delta x\to 0}\frac{1}{\pi}\sum f(x_i)x_i\frac{1}{\sqrt{1-x_i^2}}\Delta x_i$$

$$= \frac{2}{\pi}\int_{-1}^{1} f(x)x\frac{1}{\sqrt{1-x^2}}dx \qquad (5)$$

The factor 2 appears in the last expression, since summation is over only a half-cycle. For a graphical interpretation, equation 5 may be replaced by

$$B_1 = \frac{2}{\pi}\int_{-1}^{1} F_1(x)dx \qquad (6)$$

where $F_1(x)$ is a new function derived from $f(x)$ by the relation

$$F_1(x) = f(x)x\frac{1}{\sqrt{1-x^2}} \qquad (7)$$

The coefficient B_1 is proportional to the area under this new curve.

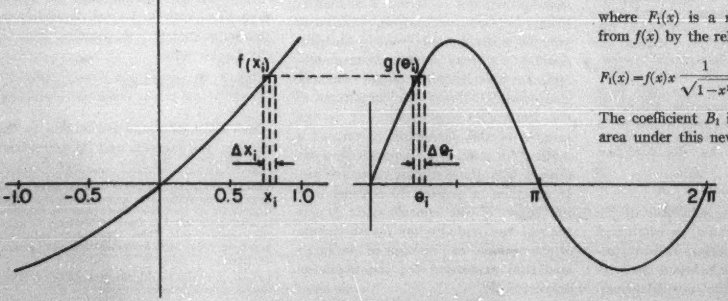


Fig. 2. Relation between variables in x and θ domains

Table I. Use of Equation 6 Illustrated

x	x_i	$f(x_i)$	$\sqrt{1-x_i^2}$	$F(x_i)$
−1.0				
	−0.9	−2.79	0.4359	5.760
−0.8				
	−0.7	−2.31	0.7141	2.264
−0.6				
	−0.5	−1.75	0.8660	1.010
−0.4				
	−0.3	−1.11	0.9539	0.349
−0.2				
	−0.1	−0.39	0.9949	0.039
0				
	0.1	0.41	0.9949	0.041
0.2				
	0.3	1.29	0.9539	0.416
0.4				
	0.5	2.25	0.8660	1.299
0.6				
	0.7	3.29	0.7141	3.225
0.8				
	0.9	4.41	0.4359	9.105
1.0				
Total				23.508

$$B_1=\frac{2}{\pi}\sum F(xi)\Delta xi.$$

$$=\frac{2}{\pi}(23.508)(0.2)=2.9931.$$

Illustration by a Numerical Evaluation

To illustrate the use of equation 6 a calculation is shown in Table I, based on the function shown in Fig. 3. For this example, a specific function

$$f(x)=4x+x^2 \qquad (8)$$

is used in order that an exact answer will be available for comparison. Ten equal intervals have been chosen for calculating an approximate value for B_1. The result obtained is 2.993. When the relation

$$x=\sin\theta$$

is substituted in equation 8 there results

$$g(\theta)=1/2+4\sin\theta-(1/2)\sin 2\theta \qquad (9)$$

and, clearly, the correct value for B_1 is 4.0. The approximate value from Table I is seen to differ considerably from the true value. However, the reason for the discrepancy is not hard to discover. It is noted the $F_1(x)$ approaches infinity at the two extremes, $x=\pm1$; hence, the areas under the curve in these neighborhoods are particularly significant, and the relatively coarse subdivision of the abscissa has not adequately emphasized these intervals.

As a means of improving the calculation, it is first observed that the function $F_1(x)$ approaches infinity near the limits, not because of $f(x)$, which remains finite, but because of the factor $\sqrt{1-x^2}$, which is in the denominator. This suggests that equation 6 may be written as

$$B_1=\frac{2}{\pi}\int_{-1}^{-a} f(x)\frac{x}{\sqrt{1-x^2}}dx+\frac{2}{\pi}\int_{-a}^{a} F_1(x)dx+\frac{2}{\pi}\int_{a}^{1} f(x)\frac{x}{\sqrt{1-x^2}}dx \qquad (10)$$

Furthermore, if $(1-a)$ is small, the first and third integrals will not be altered greatly when the variable values of $f(x)$ are replaced by constant values, selected from values in these end intervals. Thus

$$B_1=f(\xi_1)\frac{2}{\pi}\int_{-1}^{-a}\frac{x}{\sqrt{1-x^2}}dx+\frac{2}{\pi}\int_{-a}^{a}F_1(x)dx+f(\xi_2)\frac{2}{\pi}\int_{a}^{1}\frac{x}{\sqrt{1-x^2}}dx$$
$$=-f(\xi_1)\frac{2}{\pi}\sqrt{1-a^2}+\frac{2}{\pi}\int_{-a}^{a}F_1(x)dx+f(\xi_2)\frac{2}{\pi}\sqrt{1-a^2} \qquad (11)$$

where ξ_1 and ξ_2 are in the end intervals. Utilizing equation 11, an improved value for B_1 is obtained for the example of Fig. 3. First, the middle integral is evaluated using the values in Table I in the range $-0.8<x<+0.8$. Added to this are the values for the first and third terms, for $a=0.8$ and for the values of $f(x)$ taken at the mid-points of the end intervals, that is, at $\xi_1=-0.9$ and $\xi_2=+0.9$. The improved value for B_1 is 3.850, which is considerably closer than the previous calculation.

The use of a value for $f(x)$ selected from an end interval assures, in the example, that the calculation for that interval will be bounded, since $f(x)$ is bounded and, in fact, these bounds could be calculated. However, it is evident that the best choice for ξ_1 and ξ_2 is not at the mid-points of the end intervals, since the integrands in equation 10 are increasing toward ±1. There are, in fact, particular values for ξ_1 and ξ_2 which will make the first and third terms exact. If, now, it be assumed that the slope of $f(x)$ is essentially constant in the end intervals, then it is possible to determine the best choice for ξ_1 and ξ_2. The relation is derived in Appendix I, and is given by

$$\xi_1,\xi_2=\pm\left\{\frac{a}{2}+\frac{1}{2}\left[\frac{\pi/2-\sin^{-1}a}{\sqrt{1-a^2}}\right]\right\} \qquad (12)$$

This, of course, depends on the width of the end intervals. The relation is shown in Fig. 4. For the example of Fig. 3, with $a=0.8$ and $\xi=\pm0.936$, the improved calculation gives $B_1=3.954$, an error of only 1.05 per cent.

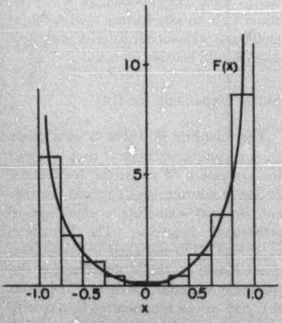


Fig. 3. Modified characteristic for evaluation of fundamental component

$$f(x)=4x+x^2$$

$$F(x)=f(x)\times\frac{1}{\sqrt{1-x^2}}$$

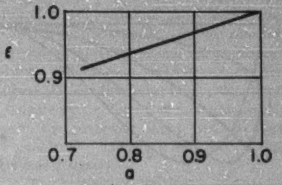


Fig. 4. Optimum co-ordinate in end interval $a<\xi<1$, applicable to fundamental component

Table II. Characteristic Functions, Equations 13

n	$\phi_n=\cos(n\theta)$	$\psi_n=\sin(n\theta)$
0	1	
1	$\sqrt{1-x^2}$ (see *)	x
2	$1-2x^2$	$2x\sqrt{1-x^2}$ (see *)
3	$(1-4x^2)\sqrt{1-x^2}$ (see *)	$3x-4x^3$
4	$1-8x^2+8x^4$	$(4x-8x^3)\sqrt{1-x^2}$ (see *)
5	$(1-12x^2+16x^4)\sqrt{1-x^2}$ (see *)	$5x-20x^3+16x^5$
6	$1-18x^2+48^4-32x^6$	$(6x-32x^3+32x^5)\sqrt{1-x^2}$ (see *)
7	$(1-24x^2+80x^4-64x^6)\sqrt{1-x^2}$ (see *)	$7x-56x^3+112x^5-64x^7$
8	$1-32x^2+160x^4-256x^6+128x^8$	$(8x-80x^3+192x^5-128x^7)\sqrt{1-x^2}$ (see *)

* For $f(x)$ double-valued, use starred functions for $f_B(x)$; unstarred functions for $f_A(x)$. For $f(x)$ single-valued, use unstarred functions.

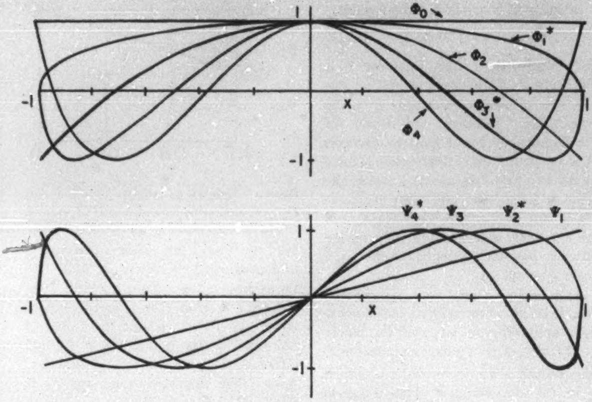


Fig. 5. Characteristic functions for sinusoidal inputs

This calculation emphasizes the importance of the end intervals. The relatively coarse subdivision of the abscissa in the middle range is not particularly important; perhaps the use of a planimeter would suffice in this range for reasonable accuracy.

It is important to note that for a numerical point solution, such as is followed here, the accuracy is inherently higher than in a point solution obtained in the θ domain with the same number of points and with equal intervals $\Delta\theta$. This is so because of the weighting factor applied to values of $f(x)$ in converting to equal intervals of Δx. Those values of $f(x)$ near the extremes, which are most important in determining the fundamental component, are given increased emphasis in the summations.

Higher Harmonics

The generalization of equations 5 and 6 for the other coefficients in equation 3 follows directly from the basic equations. These may be expressed in the general forms

$$A_n = \frac{2}{\pi}\int_{-1}^{1} f(x)\phi_n(x)\,\frac{1}{\sqrt{1-x^2}}\,dx$$
$$B_n = \frac{2}{\pi}\int_{-1}^{1} f(x)\psi_n(x)\,\frac{1}{\sqrt{1-x^2}}\,dx \qquad (13)$$

Equations 13 contain a set of functions, $\phi_n(x)$ and $\psi_n(x)$ arising from transforming cos $(n\theta)$ and sin $(n\theta)$ to the x domain. These functions are closely related to the Tschebyscheff polynomials and are derived as shown in Appendix II; examples are given in Table II. Some of the functions of Table II are illustrated in Fig. 5.

The procedure already used in evaluating the fundamental component can be adapted to any one of the coefficients. Caution may be required in the end intervals, since the special results of equations 10, 11, and 12 and Fig. 4 are based on $\psi_1(x)$ and apply only to the fundamental component. However, the general method is applicable and involves no impossible integrations; another precaution which may be necessary is mentioned in the next section.

It is clear from the procedure presented that each frequency coefficient is determined independently of all the others, so that the accuracy does not depend on the number of harmonics included, as in some methods used for distortion analysis.[2]

Single-Valued and Double-Valued Functions of X

The discussion of the foregoing example tacitly assumes that $f(x)$ is a single-valued function, that is, that $f(x)$ takes on the same values for x increasing and for x decreasing. Such a condition leads to special symmetry properties for $g(\theta)$; when the origin is taken, as in Fig. 2, the values for A_n are zero for n odd, and the values for B_n are zero for n even. The coefficients of equation 13 are expressed as being proportional to twice the integral over a half-cycle. Actually the values of ϕ_n for n odd and ψ_n for n even should be taken with opposite signs for x increasing and for x decreasing, which leads to zero value for these coefficients. Knowing this in advance, one may merely ignore these terms; an asterisk has been shown with those functions in Table II which are not to be used when $f(x)$ is single-valued.

The case of a double-valued function, such as might arise in a hysteresis curve, represents a different situation. Here each half-cycle could be calculated separately and the results added. However, an alternate procedure is suggested. This is illustrated by Fig. 6. Let

$f_1(x) = f(x)$ for x increasing

$f_2(x) = f(x)$ for x decreasing

and let two new functions be defined

$$f_A(x) = \frac{1}{2}[f_1(x) + f_2(x)]$$
$$f_B(x) = \frac{1}{2}[f_1(x) - f_2(x)] \qquad (14)$$

so that

$$f_1(x) = f_A(x) + f_B(x)$$
$$f_2(x) = f_A(x) - f_B(x) \qquad (15)$$

The function $f_A(x)$ is single-valued and its coefficients can be determined as previously described; the result will involve A_n for n odd and B_n for n even.

The function $f_B(x)$ is combined in equation 15 with $f_A(x)$ by a positive sign for x increasing and by a negative sign for x decreasing. Related to the function $f_B(x)$ is a function $g_B(\theta)$ which has symmetry about $\theta = 0$, and which has coefficients A_n only for n even and B_n only for n odd. One may evaluate the integrals for the half-cycle with x increasing and multiply by 2, again ignoring those functions whose coefficients are zero. Thus the complete result is obtained using the functions of Table II. those with an asterisk are used for $f_B(x)$ and those without an asterisk for $f_A(x)$. Equation 13 holds in all cases.

Series Expansion for f(x)

The functions in Table II as a group have a significance beyond that involved for evaluation of particular coefficients. In the θ domain, the functions cos $(n\theta)$ and sin $(n\theta)$ constitute a closed set of orthogonal functions. The functions $\phi_n(x)$ and $\psi_n(x)$ are their corresponding counterparts in the x domain (for the particular variation $x = \sin\theta$). It follows that $f(x)$ can be representeed by a combination of these functions, each one being multiplied by the corresponding coefficient of the Fourier expansion. Thus

$$f(x) = \frac{A_0}{2}[\phi_0] + \sum [A_n\phi_n(x) + B_n\psi_n(x)] \qquad (16)$$

In other words $f(x)$ can be thought of as a superposition of curves of the form shown in Fig. 5.

As an example, the particular function in equation 8 (which contains only a finite number of terms) can be written in an alternate form

$$\begin{aligned} f(x) &= 4x+x^2 \\ &= \frac{A_0}{2}[\phi_0]+B_1[\psi_1(x)]+A_2[\phi_2(x)] \\ &= \frac{A_0}{2}[1]+B_1[x]+A_2[1-2x^2] \qquad (17) \end{aligned}$$

The particular values, $A_0=1$, $B_1=4$, and $A_2=-(1/2)$, are determined by equating coefficients. This is depicted in Fig. 7, with $f(x)$ shown as the sum of three of the functions from Table II.

For an example of a double-valued function, Fig. 6 may be considered as a hysteresis loop plotted with the coordinate axes reversed from the usual form (the abscissa represents the flux and the ordinate the magnetomotive force). To determine the exciting current for a sinusoidally applied voltage, (that is, with the flux a sine wave) the curves of Fig. 6 (B) may be compared with those from Fig. 5. It is seen that the single-valued function $f_A(x)$ is essentially a combination of ψ_1 and $-\psi_3$; that is, the current contains fundamental and pronounced third-harmonic sine components. On the other hand, $f_B(x)$ is very similar in shape to ϕ_1^*, which indicates a fundamental cosine term. Thus the loop of the magnetization curve mainly contributes a shift to the fundamental component without affecting greatly the phase of the third-harmonic term.

The special significance of equation 16 lies in its relation to equation 2 which was the general form of a polynomial approximation for $f(x)$. It will be observed that there is a linear relationship between the coefficients in equations 2 and 16. Furthermore the coefficient of the highest degree term in x of equation 2 is related solely to the coefficient of the highest frequency component. One series can be converted to the other by equating coefficients of like powers.

The availability of the functions ϕ_n and ψ_n suggests two applications. The first is the determination of an appropriate polynomial representation for $f(x)$ based directly on measurements of amplitudes of the harmonic components (and measurements of phase if the function is double-valued). Substitution directly in equation 16 and rearrangement gives the values c_n, which provides the best fit in the sense that the mean square error in the θ domain is a minimum for the particular input function here considered. This is a property of the Fourier series.[8]

The second application is essentially the converse, the determination of the frequency components when $f(x)$ is specified as a polynomial. This is accomplished merely by comparing coefficients, beginning with the highest order terms. (This is shown in the example of equation 17.) The process involves only elementary operations and avoids involved trigonometric substitutions.

Extension to Inputs of Arbitrary Amplitudes

The special restriction in the preceding discussion to inputs of unit

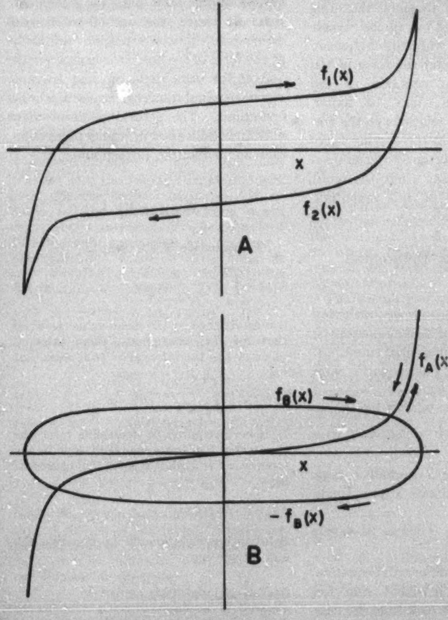


Fig. 6. A—A double-valued characteristic. B—Resolution into component functions

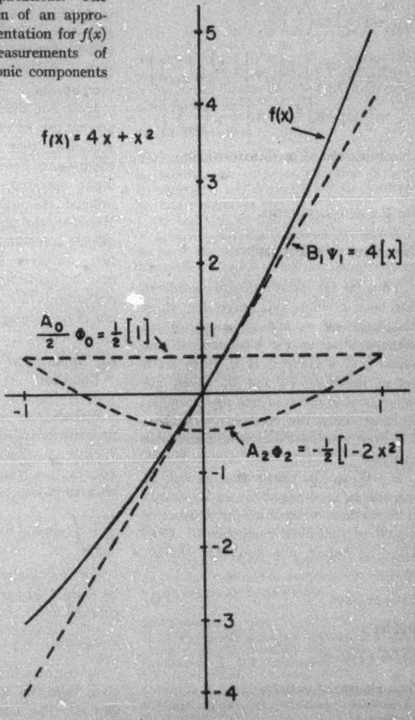


Fig. 7. A nonlinear characteristic represented as a sum of characteristic functions

amplitude may be removed with little difficulty. A more general variation is expressed by

$$x = x_0 \sin \theta \tag{18}$$

Thus far, x_0 has been assumed equal to 1. By rewriting equation 18 as

$$(x/x_0) = \sin \theta$$

which amounts to normalizing the scale of the abscissa, it is noted that the new variable (x/x_0) now has the amplitude of unity. Thus, new functions $\phi_n(x/x_0)$ and $\psi_n(x/x_0)$ can be written merely by replacing x with the new variable (x/x_0). An example will illustrate the relations. Thus, for

$$f(x) = 4x + x^2$$

let

$$x = 2 \sin \theta$$

Then

$$\begin{aligned} f(x) &= 4x_0\left(\frac{x}{x_0}\right) + x_0^2\left(\frac{x}{x_0}\right)^2 \\ &= 8\left(\frac{x}{x_0}\right) + 4\left(\frac{x}{x_0}\right)^2 \\ &= \frac{A_0}{2}[\phi_0] + B_1\left[\psi_1\left(\frac{x}{x_0}\right)\right] + A_2\left[\phi_2\left(\frac{x}{x_0}\right)\right] \\ &= \frac{A_0}{2}[1] + B_1\left[\frac{x}{x_0}\right] + A_2\left[1 - 2\left(\frac{x}{x_0}\right)^2\right] \end{aligned}$$

A comparison of coefficients yields

$$A_2 = -2 \qquad B_1 = 8 \qquad A_0 = 4$$

for the new amplitudes.

Nonsinusoidal Inputs

Thus far the input function considered has been a single sine wave, for which the functions ϕ_n and ψ_n constitute an orthogonal set in the θ domain. If the input is of a different type, there are a set of functions P_n and Q_n which are orthogonal in terms of this new variable; it is precisely the coefficients of these functions which are sought in a harmonic analysis.

It will now be shown that ϕ_n and ψ_n can still be used effectively under certain circumstances when the input consists of a sum of sinusoidal components. Consider, as before, the nonlinear characteristic

$$f(x) = 4x + x^2 \tag{19}$$

and let

$$x = \sin \theta + \sin 3\theta$$

By a change of variable

$$y = \sin \theta$$

the variation of θ can now be represented by $\psi_1(y)$. Correspondingly $\sin 3\theta$ is represented by $\psi_3(y)$. Thus the variable x may be written as

$$x = [y] + [3y - 4y^3] = 4y - 4y^3$$

By making this substitution in equation 19

$$\begin{aligned} h(y) &= f(4y - 4y^3) \\ &= 4[4y - 4y^3] + [4y - 4y^3]^2 \\ &= 16y + 16y^2 - 16y^3 - 32y^4 + 16y^6 \end{aligned}$$

This new function $h(y)$ may now be represented as a superposition of $\phi_n(y)$ and $\psi_n(y)$ exactly analogous to Equation (16). The results are

$$A_0 = 2.0$$
$$A_2 = 0.5$$
$$A_4 = -1.0$$
$$A_6 = -0.5$$
$$B_1 = 4.0$$
$$B_3 = 4.0$$

In one sense, the change of variable has resulted in what amounts to a new equivalent characteristic. That is, the original characteristic $f(x)$ with an input containing two sinusoids has been replaced by a second characteristic $h(y)$ which with a single sinusoidal input ($y = \sin \theta$) results in the same frequency components.

It is not difficult to show that this process can be carried out in the reverse sequence; that is, with the frequency components known and the form of the input specified, the equation of the original characteristic may be found. However, the phase relations of the frequency components in the driving function are very important. Extensions may be made for cases where the input components are not of unit amplitude nor of equal magnitude.

An Application to a Differential Equation

Perhaps one of the most intriguing possibilities for use of these functions is in certain nonlinear differential equations. For example, Van der Pol's equation may be written as

$$\frac{dx}{dt} + \int x\,dt = \mu\left(x - \frac{x^3}{3}\right)$$

For the case where μ is small, it may be desired to approximate x as a single sinusoid

$$x = x_0 \sin t$$

Substitution in the differential equation gives zero on the left-hand side for any x_0. The solution will hold for the fundamental and to the order of μ for higher terms, if the right-hand side contains no fundamental component. Rewriting the right-hand side as

$$\begin{aligned} \mu\left[x - \frac{x^3}{3}\right] &= \mu\left[x_0\left(\frac{x}{x_0}\right) - \frac{x_0^3}{3}\left(\frac{x}{x_0}\right)^3\right] \\ &= \mu\left[B_1\psi_1\left(\frac{x}{x_0}\right) + B_3\psi_3\left(\frac{x}{x_0}\right)\right] \\ &= \mu\left\{B_1\frac{x}{x_0} + B_3\left[3\frac{x}{x_0} - 4\left(\frac{x}{x_0}\right)^3\right]\right\} \end{aligned}$$

it follows that

$$\frac{x_0^3}{3} = 4B_3$$

$$x_0 = B_1 + 3B_3$$

so that

$$\begin{aligned} B_1 &= x_0 - \frac{x_0^3}{4} \\ &= \frac{x_0}{4}(4 - x_0^2) \end{aligned}$$

Thus, for B_1 to be equal to zero, $x_0 = 2$, the value for the limit cycle, which compares to the value obtained by other methods.[4]

General Conclusions

Unlike linear systems, the nonlinear system affords little basis for generalization. However, the use of polynomial expressions directs attention primarily at the form of the nonlinear characteristic and at the same time replaces trigonometric manipulations by simpler algebraic operations. The general point of view may find application in studies of modulation and nonlinear distortion.

Appendix I

For evaluation of the integral

$$G(a) = \int_a^1 f(x)\frac{x}{\sqrt{1 - x^2}}dx \tag{20}$$

it is established by the mean value theorem[6] that for $f(x)$ continuous, there exists a value of ξ in the interval $a < \xi < 1$, such that

$$G(a) = f(\xi)\int_a^1 \frac{x}{\sqrt{1 - x^2}}dx \tag{21}$$

It is here required to determine ξ for the case where $f(x)$ has a constant slope in the interval $a < x < 1$, that is, with $f(x)$ expressed by

$$f(x) = f(a) + k(x - a) \tag{22}$$

Substituting equation 22 in equation 20, and integrating

$$G(a) = \left[f(a) - k\frac{a}{2}\right]\sqrt{1 - a^2} + \frac{k}{2}\left[\frac{\pi}{2} - \sin^{-1} a\right] \tag{23}$$

Putting $x=\xi$ in equation 22 and substituting in equation 21, there is obtained

$$G(a)=[f(a)-ka]\sqrt{1-a^2}+k\xi\sqrt{1-a^2} \quad (24)$$

Equating the two expressions and solving for ξ

$$\xi=\frac{a}{2}+\frac{1}{2}\left[\frac{\pi/2-\sin^{-1}a}{\sqrt{1-a^2}}\right] \quad (25)$$

Appendix II. Functions of Equations 13

For the variation

$$x=\sin\theta \quad (26)$$

the following functions are defined

$$\phi_n(x)=\cos(n\theta)=\cos(n\sin^{-1}x) \quad (27)$$

$$\psi_n(x)=\sin(n\theta)=\sin(n\sin^{-1}x) \quad (28)$$

From the definitions

$$\phi_0=1$$
$$\psi_1=x$$

From the trigonometric identity

$$\sin^2(n\theta)+\cos^2(n\theta)=1 \quad (29)$$

it follows that

$$\phi_1=\cos\theta=\sqrt{1-\sin^2\theta}=\sqrt{1-x^2}$$

From the identities

$$2\sin\theta\sin n\theta=-\cos(n+1)\theta+\cos(n-1)\theta \quad (30)$$

$$2\sin\theta\cos n\theta=\sin(n+1)\theta-\sin(n-1)\theta \quad (31)$$

there may be established the recursion equations

$$2x\psi_n=-\phi_{n+1}+\phi_{n-1} \quad (32)$$

$$2x\phi_n=\psi_{n+1}-\psi_{n-1} \quad (33)$$

Thus from equation 32

$$\phi_2=\phi_0-2x\psi_1=1-2x^2$$

Using equation 29

$$\psi_2=\sqrt{1-\phi_2^2}=\sqrt{1-(1-2x^2)^2}=2x\sqrt{1-x^2}$$

All the other functions can be obtained using equations 32 and 33, and the functions thus far derived.

The functions ϕ_n and ψ_n are related to the Tschebyscheff polynomials[6] of the first and second kind, defined in a similar way but in terms of the variation $x=\cos\theta$.

References

1. Alternating Current Circuits (book), **Russell M. Kerchner, George F. Corcoran.** John Wiley & Sons, Inc., New York, N. Y., 3rd edition, 1950, p. 161.

2. Radio Engineering (book), **Frederick E. Terman.** McGraw-Hill Book Company, Inc., New York, N. Y., 2nd edition, 1937, pp. 284–75.

3. The Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications (book), **N. Wiener.** John Wiley & Sons, Inc., New York, N. Y., 1949, pp. 26–27.

4. Non Linear Mechanics (book), **N. Minorsky.** J. W. Edwards, Ann Arbor, Mich., 1947, p. 197.

5. Advanced Calculus (book), **D. V. Widder.** Prentice-Hall, Inc., New York, N. Y., 1947, pp. 34, 138.

6. Special Functions of Mathematical Physics (book), **W. Magnus, F. Oberhettinger.** Chelsea Publishing Company, New York, N. Y., 1949, p. 78.

◆

Discussion

Thomas M. Stout (P. O. Box 550, Ridgefield, Conn.): The author presents a method of harmonic analysis which is radically different from the methods generally employed. In addition to the suggested applications to nonlinear differential equations, modulation, and distortion, it should be useful for the calculation of "describing functions" employed in non-linear feedback system analysis.[1]

Extensive calculations with the proposed method will be facilitated by a tabulation of numerical values for $\phi_n(x)$, $\psi_n(x)$ and $\sqrt{1-x^2}$, and the preparation of standard forms for the various harmonics, as in Table I. This preliminary work would have been simplified if the standard Tschebyscheff polynomials had been used, since tabulated values of these functions are available[2]; however, because the interrelations are simply

$$\phi_n(x)=(-1)^{n/2}T_n(x), \text{ when } n=0, 2, 4, 6\ldots$$
$$=(-1)^{\frac{(n-1)}{2}}U_n(x), \text{ when } n=1, 3, 5, 7\ldots\ldots \quad (34)$$

$$\psi_n(x)=(-1)^{\frac{(n+2)}{2}}U_n(x), \text{ when } n=0, 2, 4, 6\ldots$$
$$=(-1)^{\frac{(n-1)}{2}}T_n(x), \text{ when } n=1, 3, 5, 7\ldots \quad (35)$$

this difficulty can be handled with a little care in the use of the tables. There may be good reasons for choosing $\phi_n(x)$ and $\psi_n(x)$ in this particular way, which are not discussed in the paper.

It may be well to emphasize again that the relations given in equations 10 through 12 apply only to $\psi_1(x)$ and the fundamental component. Anyone interested in other components must develop similar relations for them. As a start in this direction, the contribution of the end intervals in the calculation of the d-c component can be approximated by

$$-\frac{2}{\pi}f(\xi_1)\left[\frac{\pi}{2}-\sin^{-1}a\right]+\frac{2}{\pi}f(\xi_2)\left[\frac{\pi}{2}-\sin^{-1}a\right] \quad (36)$$

where ξ_1 and ξ_2 are computed from

$$\xi=\frac{\sqrt{1-a^2}}{\frac{\pi}{2}-\sin^{-1}a} \quad (37)$$

Strangely enough, these values differ only slightly from those given in Fig. 4, suggesting that a single curve might serve for both the d-c and fundamental components. It would be interesting to determine whether the same relation applies for the higher harmonics.

The accuracy of the proposed method will also bear further investigation. Although the proposed method is capable of arbitrary accuracy, it requires more work than the simple 3- and 5-point equations[3] used in distortion analysis for comparable accuracy. For example, calculations were carried out for the case

$$\begin{aligned} y&=0 && x<0.5 \\ &=x-0.5 && x>0.5 \end{aligned} \quad (38)$$

where

$$x=\sin\theta \quad (39)$$

giving the series[4]

$$y=0.1090+0.1955\sin\theta+\ldots \quad (40)$$

With the 3-point equations, the series is

$$y\cong 0.1250+0.2500\sin\theta+\ldots \quad (41)$$

and, with the five-point equations, the series is

$$y\cong 0.1143+0.1982\sin\theta+\ldots \quad (42)$$

By means of the technique described in the paper, using an interval of 0.2 and appropriate mean values in the end intervals, gives for the series

$$y\cong 0.1063+0.1915\sin\theta+\ldots \quad (43)$$

The amplitudes of equation 10 are about as accurate as those of equation 9, but considerably less computation is required to obtain equation 9.

On the basis of this and other examples, it appears that the proposed method can be used most profitably to find the higher harmonics in cases where the characteristic curve is highly nonlinear. As the author points out, the various components are determined independently and, if the necessary preliminary work had been done in advance, could be computed with roughly the same amount of effort as is required to find the fundamental component.

Although numerical results can be obtained by the proposed method, its real significance probably lies in the insights which it promotes. The importance of the extreme values of $f(x)$ and the effects of a double-valued characteristic curve are readily visualized by this approach.

In this provocative paper the author has opened a vein which will be mined extensively by future investigators. Further work on the reverse process of finding the characteristic curve by harmonic measurements and the case of nonsinusoidal inputs would be particularly welcome.

References

1. A Frequency-Response Method for Analyzing and Synthesizing Contactor Servomechanisms, **R. J. Kochenburger.** *AIEE Transactions*, vol. 69, pt, I, 1950, pp. 270–83.

2. Tables of Chebyshev Polynomials (book), National Bureau of Standards, Washington, D. C., 1952.

3. See reference 2 of the paper, p. 380.

4. *Ibid.*, pp. 21–22.

Laurel J. Lewis: The author wishes to thank Dr. Stout for discussing a number of significant points which, because of space limitations, could be only briefly mentioned in the paper.

The relation of the polynomial functions ϕ_n and ψ_n to the Tschebyscheff polynomials stated by Dr. Stout is alluded to in Appendix II. It will be noted that the unstarred functions in Table II, except for a matter of sign, are the polynomials T_n, while those which are starred, except for a matter of sign, are U_n. It was only after serious reflection that the author decided on the merits of a definition not in the standard form. It will be noted from Fig. 5 that the functions ϕ_n are all even functions having the value unity at $x=$zero; conversely the functions ψ_n are odd functions, each with a positive slope at the origin. This method of segregating the polynomials is particularly helpful to one considering the contribution of a particular function to the shape of a nonlinear characteristic $f(x)$. The fact that all the functions have either the same initial value, or positive slopes, is of considerable assistance in obtaining the correct sign associated with a particular harmonic; the Tschebyscheff polynomials are not uniform in this respect. Of course, one already thoroughly familiar with Tschebyscheff polynomials would have little difficulty using them directly, with appropriate modification in the procedure.

That the values from equation 37 and equation 12 are similar for large a is partly due to similarity in values of certain trigonometric functions for small arguments. The values, however, are not identical. To calculate to four significant figures as in Dr. Stout's example would require calculation from the equation instead of use of a curve such as that in Fig. 4. Equations for the other cases are not available, but the relations required are contained in the following two equations

$$G_n(a)=\int_a^1 f(x)\phi_n(x)\frac{xdx}{\sqrt{1-x^2}}$$

$$=f(\xi_n)\int_a^1 \phi_n(x)\frac{dx}{\sqrt{1-x^2}} \qquad (44)$$

where

$$\xi_n=\frac{\int_a^1 \phi_n(x)\frac{dx}{\sqrt{1-x^2}}}{\int_a^1 \phi_n(x)\frac{dx}{\sqrt{1-x^2}}} \qquad (45)$$

The question of accuracy is very difficult to assess. Caution must be exercised lest the one example given for comparison with the 3-point and 5-point equations be interpreted as a generalization. As a matter of interest, by a different choice for the end interval in this example, the amount of computation could have been considerably reduced and, furthermore, the answer would have been exactly correct. For this example, the nonlinear characteristic has a constant slope for $x>0.5$. Thus, by using an interval $\Delta x_t=0.5$, and $a=0.5$, the entire computation involves the end interval only. Using Dr. Stout's equations 36 and 37 for the d-c term, and equations 10 and 11 for the fundamental, the exact values are obtained. In fact, this technique was specifically devised to eliminate the errors in the end intervals in those cases where the characteristic can be approximated by a straight line near the limits. It is clear that the end interval should be selected to include the full portion of the characteristic in which the slope is essentially constant.

The 3-point equations are approximations based on the assumption that the output response contains no harmonics higher than the second; the 5-point method assumes no terms higher than the fourth. For cases where these assumptions obviously do not hold, a serious doubt immediately arises as to how accurate will be the result from these simple formulas, even for the d-c and fundamental terms. The method employed in the paper makes no such assumptions, and therefore its validity is not in doubt; the use of more points results in a closer approximation. The increased computation required in using many points for a complicated characteristic is the price required to solve a more difficult problem.

Analysis of Magnetic Amplifiers by the Use of Difference Equations

P. R. JOHANNESSEN
ASSOCIATE MEMBER AIEE

THIS paper explains and illustrates a new way of analyzing magnetic amplifiers through the use of difference equations, and it shows that this method of analysis leads to a different view of the mechanics of operation of the magnetic amplifier. It will be shown that all magnetic amplifiers operate upon the same basic principles, and that these basic principles are demonstrated by the operation of the half-wave, half-cycle-response magnetic amplifier, as is shown in Fig. 1. Furthermore, this circuit can be considered the building-block of all magnetic amplifiers.

It will be shown that the rectifiers improve the performance of the amplifier not by causing feedback, but by increasing the input impedance. As a result the term "feedback," commonly used in the magnetic amplifier literature, and generally associated with the action of the rectifiers, is found to be inconsistent with the actual behavior of the magnetic amplifier. The term "feedback" in the magnetic amplifier has therefore been redefined, and this new definition of feedback is also consistent with already accepted terminology in the electronic amplifier field, as well as in the feedback control field.

It is often believed that the time constant of the magnetic amplifier is caused by the inductance of the saturable reactors. This is incorrect, and it will be shown that the main factor giving rise to a time constant is the inherent time delay in the method of control.

Approximation of the B-H Curve

The operation of the magnetic amplifier is based on the nonlinear characteristic of the *B-H* curve, or more specifically the saturation property of the core materials. To analyze the magnetic amplifier, this nonlinearity must be taken into account one way or another. There are two basically different approaches to this problem One is to approximate the *B-H* curve by analytic expressions, for instance a power series of the form[1,2]

$$H=a_1+a_2B+a_3B^2+\ldots \qquad (1)$$

Thus, the inductance of the saturable reactor will be a nonlinear function of the current through it. With present-day core materials the transition from an unsaturated state to a saturated one is extremely sharp (see Fig. 3), and to represent the *B-H* curve with a reasonable degree of accuracy (10 per cent), terms to the 15th power or even higher must be taken into account in equation 1. This method leads to complicated nonlinear differential equations, and from an engi-

---

Paper 54-509, recommended by the AIEE Magnetic Amplifiers Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted April 23, 1954; made available for printing August 30, 1954.

P. R. JOHANNESSEN is with the Massachusetts Institute of Technology, Cambridge, Mass.

This research was supported jointly by the Department of the Army, the Department of the Navy, and the Department of the Air Force under Air Force Contract No. *AF19(122)-458* with Lincoln Laboratory, Massachusetts Institute of Technology, and executed under subcontract by the Servomechanisms Laboratory, Department of Electrical Engineering, Massachusetts Institute of Technology.

Nomenclature

a = ratio of the reset resistance to the rectifier leakage resistance referred to the control circuit
a_1, a_2 = known constants
$a(n)$ = ratio of volt-time areas, defined in Fig. 2
A = volt-time area
B = magnetic flux density
e_I, e_{II}, e_r = instantaneous reactor voltage
e_c = instantaneous control voltage
e_f = instantaneous feedback voltage
e_s = instantaneous supply voltage
e_x = instantaneous rectifier reverse voltage
e_{out} = instantaneous output voltage
E = average voltages, averaged over a half-cycle of the supply voltage. The same subscripts as for instantaneous voltages are used
f = frequency of supply voltage, cycles per second
$G(s)$ = transfer function
H = magnetic intensity
i = instantaneous current
i_{Rm} = magnetizing current corresponding to the reactor loop width referred to the output winding
K_f = feedback voltage gain
K_F = form factor
K_p = power gain
K_v = forward voltage gain
M = turns-ratio of "feedback winding" to control winding
$n-1$, n, $n+1$ = three consecutive half-cycles of the supply voltage
N = turns-ratio of output winding to control winding
R_c = resistance of control circuit
R_L = resistance of load
R_w = resistance of output winding
R_x = leakage resistance of rectifier
s = complex frequency variable
t_f = time of firing
Z_{in} = input impedance
Z_m = magnetizing impedance
α = angle of supply voltage when reactor current is equal to i_{Rm}
$\Delta\phi$ = change in flux
ϕ_f = firing angle
ϕ_q = quiescent firing angle
τ = time constant
ω = frequency of supply voltage, radians per second

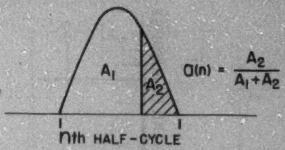


Fig. 2. Definition of the quantity $a(n)$

---

neering standpoint these equations are unsolvable.

The other method is to approximate the *B-H* curve by piecewise linear segments, Fig. 4. In the past, a transfer function of the magnetic amplifier of the form

$$G(s)=K\frac{1}{\tau s+1} \tag{2}$$

was usually assumed, and K and τ are calculated by a step-by-step procedure. By applying this method it has been possible to obtain useful expressions for the gain and the time constant in agreement with results obtained in the laboratory.[3–6]

To simplify the analysis, the approximation of Fig. 4 will be considered in this report. On the basis of this approximation, the transfer function of the magnetic amplifier will be derived by the use of difference equations.

Method of Attack

The magnetic amplifier is an average proportional amplifier. As such, the instantaneous variation of the output as a function of the input has no meaning. The operation of the magnetic amplifier, however, can be fully described by expressing the average output voltage in terms of the average input voltage. This section describes in general terms the procedure that will be followed in deriving this relationship, and it also shows that this relationship must be in the form of a difference equation.

A magnetic amplifier circuit may consist of one or more saturable reactors. Fundamentally, the operation of each reactor can be divided into two periods, the period when the reactor is unsaturated and the period when it is saturated. Output may occur in either of these two periods, depending upon what type of circuit is used. Over a given period the flux swing is proportional to the volt-time integral applied to the reactor, which again is proportional to the average reactor voltage. During any one unsaturated period which begins and ends with saturation, the total change in the flux must be zero; hence the average reactor voltage during this period is zero. This voltage is a function of the control voltage and the output voltage, and by writing it in terms of average values and equating it to zero, the equation that determines the dynamic and static behavior on an average basis has been found. This equation will express the average output voltage as a function of the average input voltage during an arbitrary period of operation. Thus, the determining equation must be in a difference form.

A variety of magnetic amplifier circuits with resistive load have been analyzed by following the foregoing procedure. The simplicity of analysis will be demonstrated by considering some typical circuits.

Series-Connected Saturable Reactor

General Considerations

The series-connected saturable reactor is shown in Fig. 5. The supply voltage e_s is a constant sinusoidal voltage of frequency ω, and in order to simplify the analysis the control voltage e_c is assumed to be a full-wave rectified sinusoidal voltage of the same frequency as e_s. Furthermore, only the case when both reactors are unsaturated for zero control voltage, usually termed normal or under-excitation, will be considered.

When the reactors are unsaturated, the impedances of both the control circuit and the output circuit are infinitely large (the magnetizing current is assumed neg-

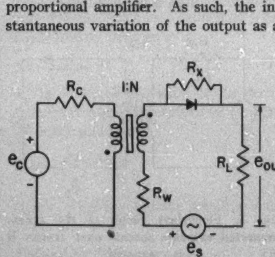


Fig. 1 (left). The basic magnetic amplifier

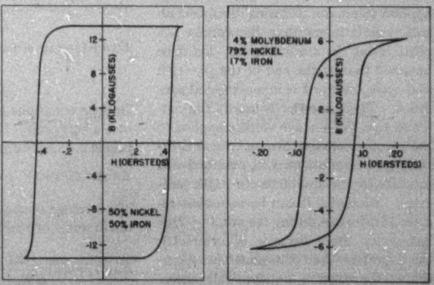


Fig. 3 (right). B-H curves for typical magnetic amplifier core materials

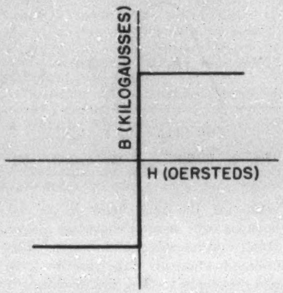


Fig. 4. Approximation of the B-H curve

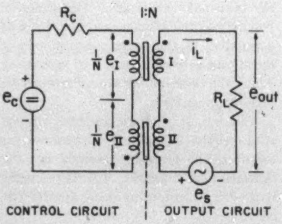


Fig. 5. The series-connected saturable reactor

ligible). Thus no currents will flow in the circuits. If any one of the reactors saturates, the load current i_L will be

$$i_L = \frac{e_s \pm Ne_c}{N^2R_c + R_L} \tag{3}$$

The polarity of the control voltage reflected into the output circuit depends upon the winding polarities of the unsaturated reactor.

If the polarities during the nth half-cycle are as shown in Fig. 5, the voltage applied to reactor II is larger than the voltage applied to reactor I. During the next half-cycle the supply voltage will reverse polarity while the polarity of e_c remains the same. As a result, the voltage applied to reactor I during the $(n+1)$th half-cycle is larger than the voltage applied during the nth half-cycle. Let it be assumed that at the end of the $(n-1)$th half-cycle, reactor I is in a saturated condition. Then, in order to satisfy the condition that the average voltage applied to the reactor between two successive saturated periods must be zero, reactor I will saturate by the end of the $(n+1)$th half-cycle. Similarly, it can be seen that reactor II saturates during the nth, $(n+2)$th half-cycles, etc. Because of symmetry the equations that determine the operation of the reactors are identical; therefore, it is necessary to consider only reactor I.

Derivation of the Determining Equation

The equation that determines the operation of the magnetic amplifier on an average basis will now be derived by equating the average reactor voltage during an arbitrary unsaturated period (as described in the section "Method of Attack") to zero.

In Fig. 6(A) is shown the volt-time area across reactor I during the nth and the $(n+1)$th half-cycles. During the unsaturated part of the nth half-cycle the reactor voltage is

$$e_I = \frac{1}{2}(e_s - Ne_c) \tag{4}$$

In the period from 2′ to 2, reactor II is saturated, and the voltage across reactor I can be determined from the equivalent circuit shown in Fig. 7. The load current becomes

$$i_L = \frac{e_s + Ne_c}{N^2R_c + R_L} \tag{5}$$

and the reactor voltage

$$e_I = e_s - R_Li_L = \frac{N^2R_ce_s - R_LNe_c}{N^2R_c + R_L} \tag{6}$$

During the $(n+1)$th half-cycle, the voltage across reactor I, from 2 to 3′ in Fig. 6(B), is

$$e_I = \frac{1}{2}(e_s + Ne_c) \tag{7}$$

To simplify the analysis, let $a(n)$ be the ratio of the volt-time area of e_s during the saturated period to the total volt-time area of e_s in the nth half-cycle. This is illustrated in Fig. 2. By the use of this quantity the average values of the reactor voltages, as shown in Fig. 6, can be written

$$E_I(n) = (1 - a(n))\frac{1}{2}(E_s - NE_c(n)) + a(n)\frac{N^2R_cE_s - R_LNE_c(n)}{N^2R_c + R_L} \tag{8}$$

$$E_I(n+1) = (1 - a(n+1))\frac{1}{2}(E_s + NE_s(n)) \tag{9}$$

In these equations E_s and E_c are averaged over a half-cycle. During the unsaturated period E_I must be zero; hence

$$E_I(n) = E_I(n+1) \tag{10}$$

The average output voltage can be written in terms of $a(n)$ as

$$E_{out}(n) = a(n)\frac{R_L}{N^2R_c + R_L}(E_s + NE_c(n)) \tag{11}$$

or

$$a(n) = \frac{E_{out}(n)}{\frac{R_L}{N^2R_c + R_L}(E_s + NE_c(n))} \tag{12}$$

By combining equations 8, 9, 10, and 12, and simplifying, the determining equation is obtained

$$E_{out}(n+1) = \frac{NR_L}{N^2R_c + R_L}(E_c(n+1) + E_c(n)) + \frac{R_L - N^2R_c}{R_L + N^2R_c}E_{out}(n) \tag{13}$$

It is interesting that the average output voltage during a given half-cycle is a function not only of the input voltage, but also of the average output voltage occurring during the previous half-cycle. The latter dependency will be defined as feedback. The reason for this definition of feedback will become more apparent in

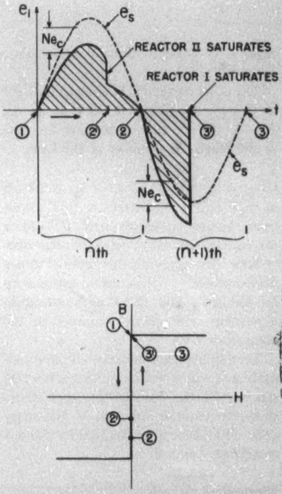


Fig. 6. A—Volt-time areas across reactor. B—Flux changes corresponding to the volt-time areas of A

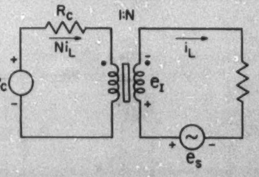


Fig. 7. Equivalent circuit of the series-connected saturable reactor when reactor II is saturated

the following analysis, and is discussed in more detail in the section "Self-Saturated Magnetic Amplifier Circuits."

Equation 13 clearly shows that the time constant and the gain do not depend upon the inductance of the control circuit or the output circuit. The gain and time constant of the series-connected saturable reactor, however, as well as for any other magnetic amplifier, can be fully explained by associating it with the phenomenon called feedback.

TRANSFER FUNCTION

The transfer function is obtained by taking the Laplace transform of equation 13 as follows

$$E_{out}(s)=\frac{NR_L}{N^2R_c+R_L}(1+e^{-\frac{\pi}{\omega}s})E_c(s)+\frac{R_L-N^2R_c}{R_L+N^2R_c}e^{-\frac{\pi}{\omega}s}E_{out}(s) \quad (14)$$

or

$$G(s)=\frac{E_{out}(s)}{E_c(s)}=\frac{\frac{NR_L}{N^2R_c+R_L}(1+e^{-\frac{\pi}{\omega}s})}{1-\frac{R_L-N^2R_c}{R_L+N^2R_c}e^{-\frac{\pi}{\omega}s}} \quad (15)$$

STABILITY CRITERIA

Equation 15 can be written in the form

$$G(s)=\frac{K_{v1}e^{-\frac{\pi}{\omega}s}}{1-K_{v1}K_{f1}e^{-\frac{\pi}{\omega}s}}+\frac{K_{v2}}{1-K_{v2}K_{f2}e^{-\frac{\pi}{\omega}s}} \quad (16)$$

where

$$K_{v1}=K_{v2}=\frac{NR_L}{N^2R_c+R_L}=K_v \quad (17)$$

$$K_{f1}=K_{f2}=\frac{R_L-N^2R_c}{NR_L}=K_f \quad (18)$$

This equation is represented by the block diagram of Fig. 8, which clearly shows that the amplifier is unstable if either one (or both) of the two parallel transmissions are unstable. The complex plane plot of the open-loop transfer function is simply a circle with center at the origin, as shown in Fig. 9. Thus, from the Nyquist stability criteria,[7] the amplifier is stable only if the feedback loop gains $K_{v1}K_{f1}$ and $K_{v2}K_{f2}$ satisfy the inequality

$$-1<K_vK_f<+1 \quad (19)$$

The feedback loop gain of the series-connected saturable reactor (see equation 16) was found to be

$$K_vK_f=\frac{R_L-N^2R_c}{R_L+N^2R_c} \quad (20)$$

Hence, it follows that the amplifier is unstable only when

$$N^2R_c=0 \text{ or } N^2R_c\rightarrow\infty \quad (21)$$

THE DETERMINING EQUATION IN DIFFERENTIAL FORM

To compare the performance of the magnetic amplifier with that of an electronic amplifier, it is convenient to write the determining equation in terms of a continuous variable. This can be done by writing it in a differential form.

A difference equation can be approximated by a differential equation, provided the sampling frequency (equivalent to the frequency of the supply voltage) is high compared to the frequency components of the input signal. As far as the magnetic amplifier is concerned, this condition is, in general, satisfied, and it is therefore legitimate to write the determining equation in a differential form.

By substituting

$$E_{out}(n+1)=(E_{out}(n+1)-E_{out}(n))+E_{out}(n) \quad (22)$$

$$E_c(n+1)=(E_c(n+1)-E_c(n))+E_c(n) \quad (23)$$

in equation 13, and dividing it by π/ω, the following expression is obtained

$$\frac{E_{out}(n+1)-E_{out}(n)}{\frac{\pi}{\omega}}+\frac{\omega}{\pi}\frac{2N^2R_c}{R_L+N^2R_c}E_{out}(n)=\frac{NR_L}{N^2R_c+R_L}\left(\frac{E_c(n+1)-E_c(n)}{\frac{\pi}{\omega}}+\frac{2\omega}{\pi}E_c(n)\right) \quad (24)$$

The corresponding differential equation is

$$\frac{d}{dt}E_{out}(t)+\frac{\omega}{\pi}\frac{2N^2R_c}{R_L+N^2R_c}E_{out}(t)=\frac{NR_L}{N^2R_c+R_L}\left[\frac{d}{dt}E_c(t)+\frac{2\omega}{\pi}E_c(t)\right] \quad (25)$$

or

$$\frac{d}{dt}E_{out}(t)+\frac{\omega}{\pi}(1-K_vK_f)E_{out}(t)=K_v\left[\frac{d}{dt}E_c(t)+\frac{2\omega}{\pi}E_c(t)\right] \quad (25)$$

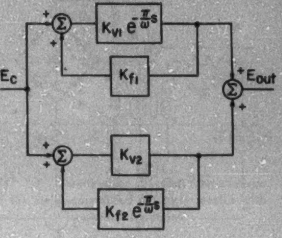
TIME CONSTANTS

By taking the Laplace transform of equation 25, an approximate transfer function is obtained as

$$G(s)=\frac{E_{out}(s)}{E_c(s)}\cong\frac{2K_v}{1-K_vK_f}\frac{\frac{\pi}{2\omega}s+1}{\frac{\frac{\pi}{\omega}}{1-K_vK_f}s+1}$$

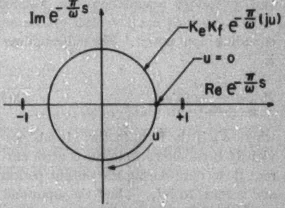
$$=\frac{R_L}{NR_c}\frac{\frac{\pi}{2\omega}s+1}{\frac{\pi}{2\omega}\left(1+\frac{R_L}{N^2R_c}\right)s+1} \quad (26)$$

From this equation the time constants will be defined as



E_c
$K_{v1}e^{-\frac{\pi}{\omega}s}$
K_{f1}
E_{out}
K_{v2}
$K_{f2}e^{-\frac{\pi}{\omega}s}$
Σ
+

Fig. 8. Block diagram of the series-connected saturable reactor



$Im\ e^{-\frac{\pi}{\omega}s}$
$-K_eK_f\ e^{-\frac{\pi}{\omega}(ju)}$
u = 0
$Re\ e^{-\frac{\pi}{\omega}s}$
-1
+1
u

Fig. 9. Complex plane plot of the function $e^{-(\pi/\omega)s}$ for $s=ju$

$$\tau_1=\frac{\pi}{2\omega} \quad (27)$$

$$\tau_2=\frac{\pi}{2\omega}\left(1+\frac{R_L}{N^2R_c}\right) \quad (28)$$

These equations show that the series-connected saturable reactor is characterized by two time constants. In previous analyses of the series-connected saturable reactor the assumption has been made that it is characterized by a single time constant only, and the time constant is defined as the time required for the step response to reach 63.3 per cent of its final value. As seen from equation 26, this assumption is valid only if

$$\tau_2\gg\tau_1 \quad (29)$$

therefore

$R_L\gg N^2R_c$

When this condition is satisfied, the transfer function (equation 26) can be written

$$G(s)\cong\frac{R_L}{NR_c}\frac{1}{\frac{\pi}{2\omega}\frac{R_L}{N^2R_c}s+1} \quad (30)$$

This is the transfer function derived by Storm,[8] and, as he points out, it is valid only when the condition of inequality 29 is satisfied.

POWER GAIN

The power gain has been defined as the ratio of the output power to the corre-

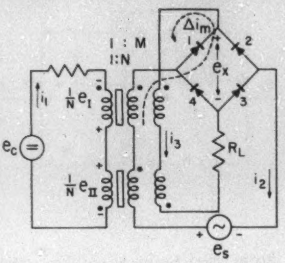


Fig. 10. The series-connected saturable reactor with external feedback

sponding input power.[3] The output power is

$$P_{out} = \frac{(E_{out})^2}{R_L}(K_F)^2 \tag{31}$$

where K_F is the form factor. It has previously been shown that the control current flows only during the output period and is equal to NI_L. Thus, the input and output form factors are identical, and the input power can be written

$$P_{in} = N\frac{E_{out}}{R_L}E_c(K_F)^2 \tag{32}$$

The power gain becomes

$$K_p = \frac{P_{out}}{P_{in}} = \frac{1}{N}\frac{E_{out}}{E_c} = \frac{R_L}{N^2R_c} \tag{33}$$

As seen from equation 33, the power gain is proportional to the steady-state voltage gain.

FIGURE OF MERIT

The figure of merit is defined as the ratio of the power gain to the time constant τ_2, thus

$$\frac{K_p}{\tau_2} = \frac{R_L}{R_L+N^2R_c}\frac{2\omega}{\pi} \tag{34}$$

It should be noted that the figure of merit is analogous to the gain-bandwidth product of the electronic amplifier.

Series-Connected Saturable Reactor with External Feedback

GENERAL CONSIDERATIONS

To improve the performance of the series-connected saturable reactor, a feedback circuit, as shown in Fig. 10, is added. The load current is rectified by a full-wave bridge and fed back through two additional feedback windings. The terms "external feedback," "feedback circuit," and "feedback windings" are the existing terminology in connection with this type of circuit. However, they are not consistent with the definition of feedback in the section "Series-Connected Saturable Reactor."

In the preceding analysis it was shown that the operation of the series-connected saturable reactor is based on feedback caused by the coupling between the two reactors. By varying R_c, the amount of feedback can be adjusted at will so as to obtain any desired gain. This, of course, will result in a corresponding change in the time constant, which is an inherent property of feedback circuits. Therefore, the improvement in the performance by adding the bridge circuit of Fig. 10 cannot be a result of additional feedback, but is caused by the switching action of the rectifiers.

DERIVATION OF THE DETERMINING EQUATION

With perfect rectifiers (infinite leakage impedance and zero forward impedance) and ideal reactors, it can be shown that the voltages across the reactors when unsaturated are undetermined. However, with very large but finite magnetizing reactance, a small but negligible magnetizing current will flow, as shown in Fig. 10, and the reactor voltages are determined uniquely. The current in the windings must follow the relations

$$i_1 = Ni_2 - Mi_3 \tag{35}$$

$$i_1 = -Ni_2 - Mi_3 \tag{36}$$

Hence

$$i_2 = 0 \tag{37}$$

The 2-valued characteristic of the rectifiers makes it necessary to assume the polarity of the rectifier bridge voltage before proceeding to determine the reactor voltages. The polarity of e_x as shown in Fig. 10 will be assumed. The current i_3, therefore, must be zero. The voltage e_x (the rectifier bridge voltage) is determined by the equations

$$Ne_x = M(e_{II} - e_I) \tag{38}$$

$$Ne_c = e_{II} - e_I \tag{39}$$

$$Ne_s = (N+M)e_{II} + (N-M)e_I \tag{40}$$

These equations may be solved as

$$e_I = \frac{1}{2}(e_s - (N+M)e_c) \tag{41}$$

$$e_{II} = \frac{1}{2}(e_s + (N-M)e_c) \tag{42}$$

$$e_x = Me_c \tag{43}$$

Thus, the polarity of the bridge voltage is as shown in Fig. 10, and consequently the current i_3 is zero. The voltages given by equations 41 and 42 are therefore the reactor voltages when the reactors are unsaturated. As seen from these equations, reactor II will saturate by the end of the half-cycle, and during this saturated period the voltage across reactor I can be found from the equivalent circuit shown in Fig. 11.

$$e_I = N\frac{(N-M)R_ce_s - R_Le_c}{R_L+(N-M)^2R_c} \tag{44}$$

The factor $a(n)$ becomes

$$a(n) = \frac{E_{out}(n)}{\dfrac{R_L}{R_L+(N-M)^2R_c}(E_s+(N-M)E_c(n))} \tag{45}$$

Because of symmetry the voltage across reactor I during the next half-cycle is given by equation 42.

From these equations the effect of the rectifier bridge circuit on the operation of the series-connected saturable reactor is apparent. The switching action of the rectifiers materially decreases the turns-ratio of the control winding to the output winding during the gating period by the factor $N/(N-M)$. Consequently the control current is reduced by the same factor. When M approaches N, the assumptions of negligible magnetizing current and infinite rectifier leakage impedance are no longer valid, because for this condition the part of the control current caused by these factors will be of the same order of mag-

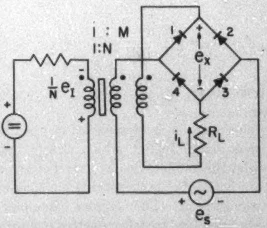


Fig. 11. Equivalent circuit of the series-connected saturable reactor with external feedback when reactor II is saturated

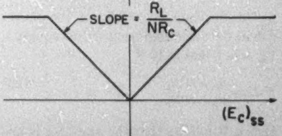


Fig. 12. Steady-state transfer characteristic of the series-connected saturable reactor

nitude as the reflected load current $(N-M)i_L$. However, if the voltage across R_c which is attributable to the magnetizing current and the rectifier leakage current is small compared to e_c, the foregoing equations are still valid. It is only when calculating the power gain that these factors are of importance.

By following the same procedure as in the section, "Series-Connected Saturable Reactor," the difference equation is found to be

$$E_{\text{out}}(n+1)=\frac{R_L}{R_L+(N-M)^2R_c}\times$$
$$((N+M)E_c(n)+(N-M)E_c(n+1))+$$
$$\frac{R_L-(N^2-M^2)R_c}{R_L+(N-M)^2R_c}E_{\text{out}}(n) \quad (46)$$

When N is equal to M, equation 46 reduces to

$$E_{\text{out}}(n+1)=2NE_c(n)+E_{\text{out}}(n) \quad (47)$$

The feedback loop gain is unity and the amplifier is unstable.

TRANSFER FUNCTION

By taking the Laplace transform of equation 46 and approximating $\epsilon^{-(\pi/\omega)s}$ by the first two terms in a Taylor series expansion (this is equivalent to the procedure followed in the preceding section), the approximate transfer function obtained is

$$G(s)\cong\frac{R_L}{(N-M)R_c}\times\frac{\frac{\pi}{2\omega}\left(1-\frac{M}{N}\right)s+1}{\frac{\pi}{2\omega}\left(\frac{N-M}{N}+\frac{R_L}{N(N-M)R_c}\right)s+1} \quad (48)$$

With reference to Fig. 8, the loop gains can be written

$$K_{v1}=\frac{(N+M)R_L}{R_L+(N-M)^2R_c} \quad (49)$$

$$K_{v2}=\frac{(N-M)R_L}{R_L+(N-M)^2R_c} \quad (50)$$

$$K_vK_f=\frac{R_L-(N^2-M^2)R_c}{R_L+(N-M)^2R_c} \quad (51)$$

When N is equal to M the transfer function reduces to

$$G(s)\cong 2N\frac{\omega}{\pi}\frac{1}{s} \quad (52)$$

As seen from this equation, for this condition $(N=M)$ the circuit behaves like a perfect integrator.

POWER GAIN AND FIGURE OF MERIT

If the magnetizing and the rectifier leakage currents are neglected, the power gain becomes

$$K_p=\frac{\frac{E_{\text{out}}^2}{R_L}}{E_c(N-M)\frac{E_{\text{out}}}{R_L}}=\frac{1}{N-M}\frac{E_{\text{out}}}{E_c}$$
$$=\frac{R_L}{(N-M)^2R_c} \quad (53)$$

and the figure of merit is

$$\frac{K_p}{\tau_2}=\frac{N}{N-M}\frac{R_L}{R_L+(N-M)^2R_c}\frac{2\omega}{\pi} \quad (54)$$

TRANSFER CHARACTERISTIC FOR NEGATIVE CONTROL VOLTAGE

The output of the series-connected saturable reactor without rectifiers is invariant to changes in the polarity of the input signal. The steady-state transfer characteristic can therefore be represented by the curve shown in Fig. 12. Adding the rectifier bridge circuit destroys this symmetry.

When the reactors are unsaturated, the voltage across the rectifier bridge is determined by the control voltage e_c. When the polarity of the control voltage is changed, a circulating current will flow in the forward direction of the rectifier, the feedback windings, and the load resistor. The reactor voltages become

$$e_{\text{I}}=\frac{1}{2}\left(e_s+\frac{NR_L}{R_L+M^2R_c}e_c\right) \quad (55)$$

$$e_{\text{II}}=\frac{1}{2}\left(e_s-\frac{NR_L}{R_L+M^2R_c}e_c\right) \quad (56)$$

From these equations it can be seen that reactor I will saturate by the end of the half-cycle, and the turns-ratio between the output winding and the control winding effectively increases by the factor $(N+M)/N$. (Compare this result with that arrived at for positive control voltage when reactor II becomes saturated.) The voltage across reactor II when reactor I saturates becomes

$$e_{\text{II}}=N\frac{(N+M)R_ce_s-R_Le_c}{R_L+(N+M)^2R_c} \quad (57)$$

The factor $a(n)$ is determined by

$$E_{\text{out}}(n)=a(n)\frac{R_L}{R_L+(N+M)^2R_c}\times$$
$$(E_s+(N+M)E_c(n))+(1-a(n))\times$$
$$\frac{MR_L}{R_L+M^2R_c}E_c(n) \quad (58)$$

By following the same procedure as that described in the section "Series-Connected Saturable Reactor," the difference equation can be found. In this report, however, no attempt will be made to derive this equation mainly because of complexity resulting from the nonlinearity which is apparent from the foregoing equations.

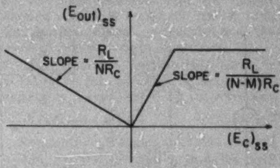


Fig. 13. Steady-state transfer characteristic of the series-connected saturable reactor with external feedback

If N is close to M and the control resistance is large compared to the load resistance, an approximate expression for the steady-state voltage gain can easily be obtained and is

$$\frac{E_{\text{out}}}{E_c}=\frac{R_L}{NR_c} \quad (59)$$

Thus, the steady-state transfer characteristic of the series-connected saturable reactor with external feedback is of the form shown in Fig. 13.

Self-Saturated Magnetic Amplifier Circuits

The group of magnetic amplifiers with self-saturation, or autoself-excitation, consists of a variety of circuit configurations. A common property of these circuits is that the load current is completely isolated from the control circuit, as was found to be the case for the series-connected saturable reactor with external feedback when the turns-ratios of the output windings and feedback windings are equal, also termed "100-per-cent feedback." As a result, the effect of self-saturation has been called "100-per-cent internal feedback."

An exact analysis of these types of circuits is extremely difficult because of the inherent nonlinearities in the rectifier characteristic and in the magnetic properties of the core materials, which are some of the most significant factors determining the performance. However, in spite of this, an idealized analysis can be justified because it will demonstrate the basic principles of operation. Also, the possibility of extending the use of difference equations to include these nonlinearities might prove feasible.

DERIVATION OF THE DETERMINING EQUATION

The circuit to be analyzed is shown in Fig. 14, and is called the self-saturated parallel magnetic amplifier. The magnetization curve is assumed to be ideal, as shown in Fig. 4, and the rectifier leakage is represented by a resistor R_x.

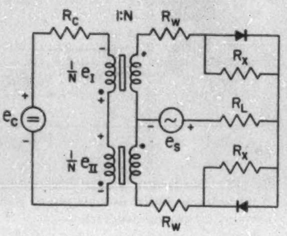


Fig. 14. The self-saturated parallel magnetic amplifier

In this circuit the output winding resistance R_w must be taken into account in the analysis. If the winding resistance is zero, the voltage across the unsaturated reactor during the gating period will be zero. This will result in an unstable condition, or a feedback loop gain of unity.

When the reactors are unsaturated, the voltages are

$$e_{\mathrm{I}}=e_s-\frac{R_x}{R_x+N^2R_c}Ne_c \tag{60}$$

$$e_{\mathrm{II}}=e_s$$

In these equations it is assumed that R_w and R_L are small compared to R_x and N^2R_c, which in general is true. As a matter of fact, this condition must be satisfied in order to obtain good performance. By the end of the half-cycle reactor II will be saturated, and the voltage across reactor I becomes

$$e_{\mathrm{I}}=\frac{N^2R_c}{R_x+N^2R_c}\frac{R_w}{R_L+R_w}e_s-\frac{R_x}{R_x+N^2R_c}e_c \tag{61}$$

and the factor $a(n)$ is determined by the equation

$$a(n)=\frac{E_{\mathrm{out}}(n)}{\dfrac{R_L}{R_L+R_w}E_s} \tag{62}$$

By following the same procedure as in the section "Series-Connected Saturable Reactor," the following difference equation is obtained

$$E_{\mathrm{out}}(n+1)=\frac{R_L}{R_L+R_w}\frac{NR_x}{R_x+N^2R_c}E_c(n)+\left(1-\frac{N^2R_c}{R_x+N^2R_c}\frac{R_w}{R_L+R_w}\right)E_{\mathrm{out}}(n) \tag{63}$$

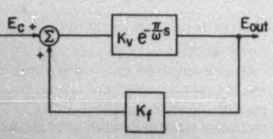


Fig. 15. Block diagram of the self-saturated parallel magnetic amplifier

The forward transmission is

$$K_v=\frac{R_L}{R_L+R_w}\frac{NR_x}{R_x+N^2R_c} \tag{64}$$

and the feedback transmission is

$$K_f=\frac{1}{N}\left(\frac{R_L+R_w}{R_L}+\frac{N^2R_c}{R_x}\right) \tag{65}$$

The approximate transfer function becomes

$$G(s)\cong\frac{R_LR_x}{R_wNR_c}\frac{1}{\dfrac{\pi}{\omega}\left(1+\dfrac{R_L}{R_w}\right)\left(1+\dfrac{R_x}{N^2R_c}\right)s+1} \tag{66}$$

FEEDBACK AND INPUT IMPEDANCE

It has been shown that the transfer function of magnetic amplifiers can be divided into two parts: a forward transmission and a feedback transmission, as illustrated in Figs. 8 and 15. The forward transmission (see equations 17, 49, 50, and 64) is equivalent to the voltage across the reactor caused by the control voltage and multiplied by the ratio $R_L/(R_L+R_w)$. It is well known that the forward transmission uniquely determines the gain-bandwidth product (the "gain" in this product refers to the voltage gain) of an amplifier, the amount of feedback is immaterial. The figure of merit of magnetic amplifiers has been defined as the ratio of the power-gain to the time constant. The time constant is inversely proportional to the bandwidth, and consequently the only difference between the gain-bandwidth product and the figure of merit for a magnetic amplifier is the different definitions of gain used. To demonstrate that the definition of feedback presented in this paper is consistent with already accepted terminology in the electronic amplifier field, as well as in the feedback control field, it is desirable to show at this time that the gain-bandwidth product of the magnetic amplifier is invarient with the amount of feedback.

The gain-bandwidth product is proportional to the ratio of the voltage gain to the time constant. In Fig. 8 is shown a general block diagram which is valid for all magnetic amplifiers considered in this report. The general transfer function becomes

$$G(s)=\frac{K_{v1}e^{-\frac{\pi}{\omega}s}}{1-K_{v1}K_{f1}e^{-\frac{\pi}{\omega}s}}+\frac{K_{v2}}{1-K_{v2}K_{f2}e^{-\frac{\pi}{\omega}s}} \tag{67}$$

but $K_{v1}K_{f1}=K_{v2}K_{f2}=K_vK_f$ (see equations 17, 18, and 51). Thus

$$G(s)=\frac{K_{v1}e^{-\frac{\pi}{\omega}s}+K_{v2}}{1-K_vK_fe^{-\frac{\pi}{\omega}s}}=\frac{K_{v1}+K_{v2}e^{\frac{\pi}{\omega}s}}{e^{\frac{\pi}{\omega}s}-K_vK_f} \tag{68}$$

By approximating $e^{\frac{\pi}{\omega}s}$, by the first two terms in a Taylor series expansion, an approximate transfer function is obtained

$$G(s)\cong\frac{K_{v1}+K_{v2}\left(1+\dfrac{\pi}{\omega}s\right)}{1+\dfrac{\pi}{\omega}s-K_vK_f}$$

$$=\frac{K_{v1}+K_{v2}}{1-K_vK_f}\frac{\dfrac{K_{v2}}{K_{v1}+K_{v2}}\dfrac{\pi}{\omega}s+1}{\dfrac{\dfrac{\pi}{\omega}}{1-K_vK_f}s+1} \tag{69}$$

From equation 69 the voltage gain and the time constant are found to be

$$K=\frac{K_{v1}+K_{v2}}{1-K_vK_f} \tag{70}$$

$$\tau=\frac{\dfrac{\pi}{\omega}}{1-K_vK_f} \tag{71}$$

and the ratio of the voltage gain to the time constant

$$\frac{K}{\tau}=\frac{K_{v1}+K_{v2}}{\dfrac{\pi}{\omega}}=(K_{v1}+K_{v2})2f \tag{72}$$

Equation 72 shows that the gain-bandwidth product of the magnetic amplifier is a function only of the forward gain and the frequency of the supply voltage.

From the foregoing analysis the importance of the forward transmission is apparent. In the following analysis the forward gain will be written in terms of the input impedance, and it will be shown that the addition of the rectifiers increases the input impedance. This, in turn, allows a higher turns-ratio N and hence an increased forward gain.

The forward voltage gain can be written in the form

$$K_v=N'\frac{Z_{\mathrm{in}}}{R_c+Z_{\mathrm{in}}}\frac{R_L}{R_L+R_w} \tag{73}$$

For the three amplifiers analyzed, N' and the input impedance Z_{in} are as follows:

1. Series-connected saturable reactor

$$N'=N,\ Z_{\mathrm{in}}=\frac{1}{N^2}R_L$$

2. Series-connected saturable reactor with external feedback

K_{v1}: $N'=N+M$

K_{v2}: $N'=N-M$

$$Z_{\mathrm{in}}=\frac{1}{(N-M)^2}R_L$$

3. Self-saturated parallel magnetic amplifiers

$$N'=N,\ Z_{\mathrm{in}}=\frac{1}{N^2}R_x$$

If the rectifier leakage impedance is taken into account in the analysis of the series-connected saturable reactor with external feedback, the input impedance becomes

$$Z_{\text{in}}=\frac{\left(\frac{1}{(N-M)^2}R_L\right)\left(\frac{1}{N^2}\frac{R_x}{2}\right)}{\frac{1}{(N-M)^2}R_L+\frac{1}{N^2}\frac{R_x}{2}} \tag{74}$$

When N is equal to M equation 74 reduces to

$$Z_{\text{in}}=\frac{1}{N^2}\frac{R_x}{2} \tag{75}$$

Equation 74 shows that the addition of the rectifiers increases the input impedance. This can also be seen by comparing the input impedance of the series-connected saturable reactor with the input impedance of the self-saturated parallel magnetic amplifier.

For a given control source with a fixed internal impedance maximum voltage gain of the forward transmission is obtained when the impedance of the control source is matched to the input impedance of the magnetic amplifier (equivalent to maximum power transfer from control source to magnetic amplifier). Thus, in order to obtain maximum voltage gain of the forward transmission, or maximum gain-bandwidth product, the following condition must be satisfied

$$R_c=Z_{\text{in}} \tag{76}$$

The condition represented by equation 76 can be fulfilled by adjusting the turns-ratio N. Hence, if the input impedance is increased, the voltage gain of the forward transmission can be increased by adjusting the turns-ratio N so that equation 76 is satisfied.

The maximum value of the ratio of the voltage gain to the time constant for the three amplifiers analyzed becomes as follows:

1. Series-connected saturable reactor

$$\left|\frac{E_{\text{out}}/E_c}{\tau}\right|_{\max}=\sqrt{\frac{R_L}{R_c}(2f)} \tag{77}$$

2. Series-connected saturable reactor with external feedback when $N=M$

$$\left|\frac{E_{\text{out}}/E_c}{\tau}\right|_{\max}=\sqrt{\frac{R_x}{2R_c}(2f)} \tag{78}$$

3. Self-saturated parallel magnetic amplifier

$$\left|\frac{E_{\text{out}}/E_c}{\tau}\right|_{\max}=\frac{1}{2}\sqrt{\frac{R_x}{R_c}\frac{R_L}{R_L+R_w}(2f)} \tag{79}$$

To demonstrate further the usefulness of the method of analysis presented, it will be shown in the following section that the forward transmission is identical to the transfer function of the basic magnetic amplifier shown in Fig. 1.

Basic Magnetic Amplifier

It was pointed out at the start of the paper that the basic magnetic amplifier shown in Fig. 1 is the fundamental circuit, or the building-block of all magnetic amplifiers. This circuit will therefore be analyzed in more detail than the other circuits.

The basic magnetic amplifier circuit is divided into two parts: the control circuit and the output circuit. The control circuit consists of the control element (the control voltage source e_c, in series with a resistor R_c) and the control winding of the saturable reactor. The output circuit consists of a fixed sinusoidal varying voltage source e_s, the output winding of the saturable reactor, a rectifier, and the load resistor R_L.

The operation of the basic magnetic amplifier is characterized by three distinct modes of operation distributed in time. The first mode is called the control period; the second and third modes are called the output period. The output period is subdivided into two parts, (corresponding to the second and the third modes), the unsaturated and the saturated period; the latter is also called the gating peroid.

The control period is that part of the operating cycle during which the sum of the voltages in the output circuit is in such a direction as to tend to drive current in the reverse direction of the rectifier. The output period is that part of the operating cycle during which the sum of the voltages in the output circuit is in such a direction as to drive current in the forward direction of the rectifier.

The impedances seen by the control element during the three modes of operation are:

First mode

$$Z_{\text{in}_1}=\frac{1}{N^2}\frac{Z_m(R_x+R_L+R_w)}{Z_m+(R_x+R_L+R_w)} \tag{80}$$

Second mode

$$Z_{\text{in}_2}=\frac{1}{N^2}\frac{Z_m(R_L+R_w)}{Z_m+R_L+R_w} \tag{81}$$

Third mode

$$Z_{\text{in}_3}=\frac{1}{N^2}Z_m \text{ saturated} \tag{82}$$

In these equations Z_m is the magnetizing impedance of the reactor.

To simplify the analysis the following assumptions are made:

1. The voltage source e_s is at every instant of time greater than the voltage across the saturable reactor.
2. The load resistance and the output winding resistance are much smaller than the rectifier leakage impedance and the reflected control source impedance.
3. The saturable reactor B-H curve is as shown in Fig. 4. The rectifier leakage impedance is represented by the resistance R_x.

Assumption 1 states that the control periods are the negative half-cycles, and the output periods are the positive half-cycles of the fixed voltage source e_s. Assumption 2 is consistent with the desired characteristics of the amplifier, as will be shown later.

MECHANICS OF OPERATION OF BASIC MAGNETIC AMPLIFIER

First, the operation of the amplifier when the control voltage is equal to zero will be determined. As a starting point, the assumption will be made that at the end of the output period, or gating period, the reactor is saturated. The voltage applied to the reactor during the control period is in such a direction as to decrease the flux, and the amount of flux change is determined by the equation

$$\Delta\phi_R=K\int_C e_{R1}dt \tag{83}$$

where K is a constant and the integration is performed over the control period. The voltage applied to the reactor in the control period is

$$e_{R1}=\frac{N^2R_c}{N^2R_c+R_x}e_s \tag{84}$$

where

$$e_s=E_{s_{\text{peak}}}\sin\omega t \tag{85}$$

By substituting equations 84 and 85 in equation 83

$$\Delta\phi_R=K\frac{N^2R_c}{N^2R_c+R_x}E_s\frac{\pi}{\omega} \tag{86}$$

E_s is the average value of the supply voltage magnitude during a half-cycle.

During the second mode of operation, the unsaturated part of the output period, the voltage applied to the reactor reverses polarity. When the volt-time integral applied to the reactor during the second mode of operation is equal to the volt-time integral applied during the first mode of operation (the control period), the reactor reaches saturation and the third mode of operation (the gating period), starts. The transition from the second to the third mode of operation is called the time of firing and is determined by the equation

$$K\frac{N^2R_c}{N^2R_c+R_x}E_s\frac{\pi}{\omega}=K\int_{\frac{\pi}{\omega}}^{t_f}e_{R2}dt \tag{87}$$

where

t_f = time of firing
e_{R2} = voltage across the reactor during the second mode of operation

The voltage e_{R2} can be written

$$e_{R2}=\frac{N^2R_c}{N^2R_c+R_L}e_s\cong E_{s_{\text{peak}}}\sin\omega t \tag{88}$$

Thus

$$t_f=\frac{1}{\omega}\cos^{-1}\left(1-\frac{2N^2R_c}{N^2R_c+R_x}\right)+\frac{\pi}{\omega} \tag{89}$$

Fig. 16 shows the voltages across the reactor during the three modes of operation. The shaded volt-time areas in Fig. 16 are proportional to the changes in flux. The time of firing t_f is determined so that the two areas are equal. The time of firing can also be expressed in terms of the firing angle, defined as

$$\phi_f=\cos^{-1}\left(1-\frac{2N^2R_c}{N^2R_c+R_x}\right) \tag{90}$$

During the third mode of operation which is the gating period, the voltage e_s appears across the load R_L and the resistance of the output winding R_w.

This analysis started with the assumption that the reactor was saturated at the end of the output period. It has been shown that this assumption is valid; hence, the next cycle of operation will be identical to the one just analyzed. A steady-state condition therefore exists.

Now let a voltage e_c be applied in the control circuit. During the control period, the voltage across the reactor attributed to the voltage e_c is

$$\Delta e_{R1}=\frac{R_x}{N^2R_c+R_x}Ne_c \tag{91}$$

and the change in the flux caused by this voltage is

$$\Delta\phi_c=K\int_0^{\frac{\pi}{\omega}}\frac{R_x}{N^2R_c+R_x}Ne_c dt \tag{92}$$

$$=K\frac{R_x}{N^2R_c+R_x}NE_c'\frac{\pi}{\omega}$$

where E_c' is the average control voltage during the control period. During the second mode of operation, the voltage appearing across the reactor attributable to e_c is

$$\Delta e_{R2}=\frac{R_L}{N^2R_c+R_L}Ne_c\cong 0 \tag{93}$$

The change in the flux during the control period must be matched by a change during the output period. Equation 93 states that a voltage appearing in the control circuit during the output period does not cause any change in the voltage across the reactor, hence there is no change in the flux. A voltage appearing in the control circuit during the output period can therefore be neglected, and the increase in the output volt-time integral attributable to e_c is

$$\Delta\int e_{\text{out}}dt=\frac{R_x}{N^2R_c+R_x}\frac{R_L}{R_L+R_w}NE_c'\frac{\pi}{\omega} \tag{94}$$

This is shown graphically in Fig. 17.

VOLTAGE GAIN

Analogous to the electronic amplifier, the operating condition that exists when the control voltage is zero can be termed "the quiescent operating condition." The output voltage is then defined as the voltage appearing across the load attributable to the control voltage. It is clear from the nature of the magnetic amplifier output that a proportionality constant has no meaning in regard to instantaneous variations of the input, but a proportionality factor can be defined on an average basis.

Because the average output voltage is a function only of the control voltage that exists during the control period, the average control voltage is defined considering it to be zero during the output period. Thus, the relation between average output and input voltages can be written

$$E_{\text{out}}=\frac{NR_x}{N^2R_c+R_x}\frac{R_L}{R_L+R_w}E_c \tag{95}$$

where E_c is the average control voltage.

The average voltage gain is

$$K_v=\frac{E_{\text{out}}}{E_c}=\frac{NR_x}{N^2R_c+R_x}\frac{R_L}{R_L+R_w} \tag{96}$$

For a given R_x and R_c, maximum voltage gain is obtained when

$$N^2R_c=R_x \tag{97}$$

giving

$$(K_v)_{\max}=\frac{N}{2}\frac{R_L}{R_L+R_w}=\frac{1}{2}\sqrt{\frac{R_x}{R_c}}\frac{R_L}{R_L+R_w} \tag{98}$$

TRANSFER FUNCTION

The average output voltage in the $(n+1)$th half-cycle is

$$E_{\text{out}}(n+1)=\frac{NR_x}{N^2R_c+R_x}\frac{R_L}{R_L+R_w}E_c(n) \tag{99}$$

where

$$n=\frac{\omega}{\pi}t,\ \text{for } t=0,\frac{\pi}{\omega},2\frac{\pi}{\omega},3\frac{\pi}{\omega},\ldots \tag{100}$$

By taking the Laplace transform of equation 99 the transfer function of the amplifier is obtained.

$$G(s)=\frac{E_{\text{out}}(s)}{E_c(s)}=\frac{NR_x}{N^2R_c+R_x}\times$$

$$\frac{R_L}{R_L+R_w}\epsilon^{-\frac{\pi}{\omega}s}=K_v\epsilon^{-\frac{\pi}{\omega}s} \tag{101}$$

If the modulation frequency of the control voltage is small compared to the carrier frequency (the frequency of the supply voltage), the difference equation 99 can be approximated by the following differential equation

$$\frac{d}{dt}E_{\text{out}}(t)+\frac{\omega}{\pi}E_{\text{out}}(t)=\frac{\omega}{\pi}K_vE_c(t) \tag{102}$$

The corresponding approximate transfer function is

$$G(s)\cong\frac{K_v}{\frac{\pi}{\omega}s+1} \tag{103}$$

From equation 103 an effective time constant can be defined as

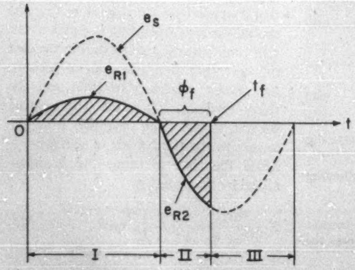


Fig. 16 (left). Quiescent volt-time areas in the basic magnetic amplifier

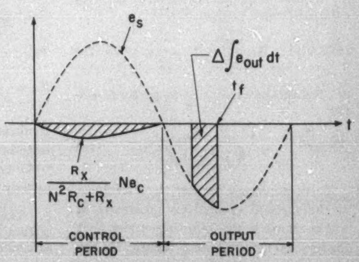
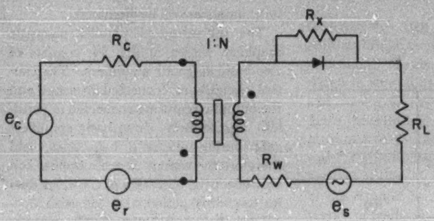


Fig. 17 (right). Volt-time areas in the basic magnetic amplifier caused by the control voltage



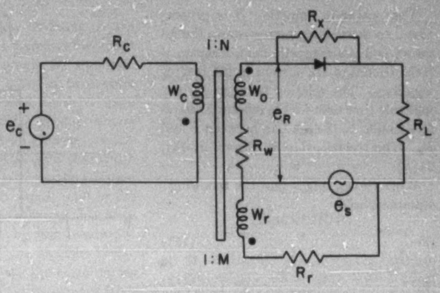


Fig. 18 (above). Reset voltage source in the control winding

Fig. 19 (right). The basic magnetic amplifier with reset winding

$$\tau_e=\frac{\pi}{\omega} \qquad (104)$$

Methods of Setting the Quiescent Firing Angle

Maximum voltage gain is obtained when the control source impedance R_c is equal to the reflected input impedance $\frac{1}{N^2}R_x$. For this condition the quiescent firing angle is

$$\phi_q=\cos^{-1}\left(1-\frac{2N^2R_c}{2N^2R_c}\right)=90 \text{ degrees} \qquad (105)$$

In many magnetic amplifier circuits a quiescent firing angle of 150 degrees or less will result in a large quiescent current exceeding the rating of both reactor windings and rectifiers. It is therefore of importance to be able to adjust the quiescent firing angle without decreasing the performance. Two methods to achieve this desired characteristic are as follows:

1. Reset the voltage source in the control circuit: One method of adjusting the quiescent firing angle is to connect a voltage source, usually called the reset voltage, in series with the control source, as shown in Fig. 18. It has previously been shown that the maximum voltage gain is a function only of the load resistor R_L, the output winding resistance R_w, and the leakage impedance of the rectifier when considering an ideal reactor. This method therefore does not decrease the voltage gain. It is interesting to note the analogy between the bias voltage in the electronic amplifier and the reset voltage in the magnetic amplifier.
2. Reset the winding: In Fig. 19 the use of a reset winding to adjust the quiescent firing angle is shown. The input impedance during the control period is

$$R_{\text{in}}=\frac{\frac{1}{N^2}R_x\frac{1}{M^2}R_r}{\frac{1}{N^2}R_x+\frac{1}{M^2}R_r} \qquad (106)$$

For simplicity, let

$$\frac{1}{M^2}R_r=a\frac{1}{N^2}R_x \qquad (107)$$

Substituting equation 107 into equation 106 gives

$$R_{\text{in}}=\frac{a}{1+a}\,\frac{1}{N^2}R_x \qquad (108)$$

To obtain maximum voltage gain the following equation must be satisfied

$$R_c=\frac{a}{1+a}\,\frac{1}{N^2}R_x \qquad (109)$$

and for this condition the voltage appearing across the gating winding when the control voltage is zero is

$$e_{R1}=\frac{\frac{N}{M}+a}{2(1+a)}e_s \qquad (110)$$

If a quiescent firing angle of 180 degrees is desired, the voltage across the reactor during the control period must be equal to the voltage across the reactor during the output period, giving

$$e_{R1}=e_s \qquad (111)$$

Under this condition

$$\frac{\frac{N}{M}+a}{2(1+a)}=1 \qquad (112)$$

Therefore

$$\frac{N}{M}=a+2 \qquad (113)$$

From equations 98 and 109, the voltage gain is

$$(K_v)_{\max}=\frac{N}{2}\,\frac{R_L}{R_L+R_w}$$

$$=\frac{1}{2}\sqrt{\frac{a}{1+a}}\sqrt{\frac{R_x}{R_c}}\,\frac{R_L}{R_L+R_w} \qquad (114)$$

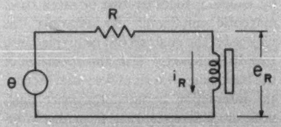


Fig. 20. Equivalent circuit of the basic magnetic amplifier

It is apparent that this method of reset will decrease the voltage gain. However, by making a sufficiently large the decrease in voltage gain can be made negligible. This necessitates a corresponding increase in the turns-ratio of the gating winding to the reset winding, as shown by equation 113.

Effect of Rectangular Hysteresis Loop

The incremental gain of the basic magnetic amplifier when a rectangular hysteresis loop is considered will now be derived. There are many other effects caused by a rectangular hysteresis loop, such as unblocking of the rectifiers during the control period, and limitation in the operating range. These effects, however, are not part of the analysis and are therefore not discussed in this paper.

By the use of Thevenin's theorem, the circuits in Figs. 18 and 19 can be reduced to the circuit shown in Fig. 20. When a voltage source in series with the control source is used to obtain reset, the circuit parameters of Fig. 20 are

$$R=\frac{N^2R_cR_x}{N^2R_c+R_x} \qquad (115)$$

$$e=\frac{N^2R_c}{N^2R_c+R_x}e_s+\frac{R_x}{N^2R_c+R_x}N(e_c+e_r) \qquad (116)$$

If instead a reset winding is used then

$$R=\frac{N^2R_cR_x'}{N^2R_c+R_x'} \qquad (117)$$

$$e=\frac{N^2R_c}{N^2R_c+R_x'}e_s'+\frac{R_x'}{R_x'+N^2R_c}Ne_c \qquad (118)$$

where

$$R_x'=\frac{\left(\frac{N}{M}\right)^2R_rR_x}{\left(\frac{N}{M}\right)^2R_r+R_x} \qquad (119)$$

$$e_s'=\frac{R_x+\frac{N}{M}R_r}{R_x+\left(\frac{N}{M}\right)^2R_r}\,\frac{N}{M}e_s \qquad (120)$$

The magnetic induction B is proportional to the volt-time area across the reactor, and the magnetic field intensity is proportional to the current in the reactor windings. The relation between current and volt-time area for the reactor is therefore similar to the relation between B and H. This relationship is shown in Fig. 21(A).

When $-i_{Rm} > i_R > i_{Rm}$, the following equations hold

$$i_R = \frac{e}{R} \tag{121}$$

$$e_R = 0 \tag{122}$$

When $i_R = i_{Rm}$

$$e_R = e - Ri_{Rm} \tag{123}$$

For simplicity, let

$$e_s = E_{s_{\text{peak}}} \sin \omega t \tag{124}$$

$$e_r = E_{r_{\text{peak}}} \sin \omega t \tag{125}$$

$$e_c = E_{c_{\text{peak}}} \sin \omega t \tag{126}$$

$$e = E_{\text{peak}} \sin \omega t \tag{127}$$

With reference to Fig. 21(B) the volt-time area across the reactor during the control period can be written

$$A = \frac{E_{\text{peak}}}{\omega} [2 \cos \alpha - (\pi - 2\alpha) \sin \alpha] \tag{128}$$

where

$$\sin \alpha = \frac{Ri_{Rm}}{E_{\text{peak}}} \tag{129}$$

The average voltage across the reactor becomes

$$E_R = \frac{1}{2} E \, [2 \cos \alpha - (\pi - 2\alpha) \sin \alpha] \tag{130}$$

where E = average value of e over a half-cycle. By taking the derivative of E_R in equation 130 with respect to E_c (the average control voltage over a half-cycle) and multiplying it by the factor $R_L/(R_L + R_w)$ the incremental voltage gain will be obtained.

$$\frac{dE_{\text{out}}}{dE_c} = \frac{NR_x'}{R_x' + N^2 R_c} \frac{R_L}{R_L + R_w} \cos \alpha \tag{131}$$

where

$$\cos \alpha = \sqrt{1 - \left(\frac{Ri_{Rm}}{E_{\text{peak}}}\right)^2} \tag{132}$$

By recognizing that R'_x in equation 132 is equal to $N^2 Z_{\text{in}}$, equation 131 can be written

$$\frac{dE_{\text{out}}}{dE_c} = N \frac{Z_{\text{in}}}{R_c + Z_{\text{in}}} \frac{R_L}{R_L + R_w} \cos \alpha \tag{133}$$

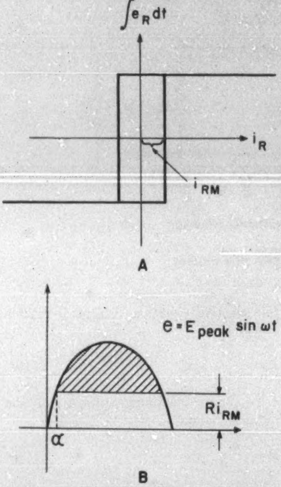


Fig. 21. A—Volt-time area as a function of current for core materials with rectangular hysteresis loop. B—The voltage across the reactor for rectangular hysteresis loop

It is interesting to compare equation 133 with equation 73. If the loop-width is zero the cos α term in equation 133 is unity and the two equations are identical. Thus, it has been shown that the forward transmission of the four amplifiers analyzed is of the same form. Furthermore, for all self-saturated magnetic amplifiers the forward transmission is identical to the forward transmission of the basic magnetic amplifier. The only basic difference between these amplifiers is the presence of feedback.

In the basic magnetic amplifier no feedback is present. That this is true can also be seen intuitively because control and output occur in different half-cycles. Consequently, there can be no interaction between control and output, and no feedback can exist.

BASIC MAGNETIC AMPLIFIER AS BUILDING BLOCK

Multicore magnetic amplifiers of the self-saturating types can be constructed by coupling together two or more basic magnetic amplifiers. By adopting this viewpoint the analysis can be divided into two distinct parts: first, the operation of the basic magnetic amplifier; and second, the interaction between the basic magnetic amplifiers resulting from this interconnection. To demonstrate this method of analysis the self-saturated parallel magnetic amplifier will be discussed.

The self-saturated parallel magnetic amplifier, shown in Fig. 14, consists of two basic magnetic amplifiers. The control circuits are connected in series and the output circuits are connected in parallel. Furthermore, these basic amplifiers gate during alternate half-cycles, and therefore the output of one amplifier can only effect the operation of the other during its control period, and vice versa.

During a given half-cycle the average output voltage is a function of the control and the output voltage during the previous half-cycle. The voltage across the reactor caused by the control voltage is

$$E_{R1} = \frac{NR_x}{N^2 R_c + R_x} E_c \tag{134}$$

and the output voltage during the next half-cycle resulting from E_{R1} is

$$\begin{aligned} E_{\text{out 1}}(n+1) &= \frac{R_L}{R_L + R_w} E_{R1}(n) \\ &= \frac{NR_x}{N^2 R_c + R_x} \frac{R_L}{R_L + R_w} E_c(n) \end{aligned} \tag{135}$$

Equation 135 represents the forward transmission of the magnetic amplifier, and is identical to the transmission of the basic magnetic amplifier.

The effect on the operation of the amplifier caused by the interconnection will now be considered. This problem will be divided into two parts, the interaction caused by the control circuits, and the interaction caused by the output circuits.

As seen from Fig. 14 the control circuits of the two basic amplifiers are connected in series. Thus, any voltages induced in one of the reactors is in series with the control voltage. As far as the other reactor is concerned, this induced voltage will appear as part of the control voltage. When one of the reactors saturates the voltage across it becomes zero, and the supply voltage e_s will appear across the load resistance R_L and the output winding resistance R_w. As far as the control circuit is concerned, this phenomena is equivalent to inserting a voltage in the control circuit equal in magnitude and with opposite polarity to the voltage that originally appeared across the reactor when the control voltage was zero. This equivalent voltage can be written in terms of the average output voltage as

$$E_{eq1} = \frac{1}{N} \frac{R_L + R_w}{R_L} E_{\text{out}} \tag{136}$$

The average voltage expressed by equation 136 is in series with the control voltage, and the output voltage during the next half-cycle caused by E_{eq1} can be written

$$E_{\text{out }2}(n+1)=\frac{NR_x}{N^2R_c+R_x}\,\frac{R_L}{R_L+R_w}\times\frac{1}{N}\,\frac{R_L+R_w}{R_L}\,E_{\text{out}}(n)$$

$$=\frac{R_x}{N^2R_c+R_x}\,E_{\text{out}}(n) \qquad (137)$$

The next step is to determine the effect on the operation of the amplifier caused by the interaction of the output circuits. From Fig. 14 it is seen that the load R_L is common to both circuits. The output voltage will therefore effect the reset of the unsaturated reactor through the rectifier leakage impedance. The polarity of the output voltage opposes the supply voltage and thus tends to oppose the amount of reset. As far as the transfer characteristic is concerned, a voltage in the output circuit that controls the flux level of the unsaturated reactor can be transferred into the control circuit by multiplying it by the turns-ratio and the ratio of the control circuit impedance to the output circuit impedance. Thus

$$E_{eq2}=\frac{1}{N}\,\frac{N^2R_c}{R_x}\,E_{\text{out}} \qquad (138)$$

The average voltage expressed by equation 138 can be considered in series with the control voltage, and the output voltage during the next half-cycle caused by E_{eq2} can be written

$$E_{\text{out }3}(n+1)=\frac{NR_x}{N^2R_c+R_x}\,\frac{R_L}{R_L+R_w}\times\frac{1}{N}\,\frac{N^2R_c}{R_x}\,E_{\text{out}}(n)$$

$$=\frac{N^2R_c}{N^2R_c+R_x}\,\frac{R_L}{R_L+R_w}\,E_{\text{out}}(n) \qquad (139)$$

The two voltages expressed by equations 136 and 138 are proportional to the output voltage. The feedback gain is defined as the portion of the output voltage that is fed back into the control circuit. Thus, the feedback gain becomes

$$K_f=\frac{1}{N}\left(\frac{R_L+R_w}{R_L}+\frac{N^2R_c}{R_x}\right) \qquad (140)$$

The output voltage is obtained by adding equations 135, 137, and 139.

$$E_{\text{out}}(n+1)=\frac{NR_x}{R_x+N^2R_c}\,\frac{R_L}{R_L+R_w}\,E_c(n)+\left(1-\frac{N^2R_c}{R_x+N^2R_c}\,\frac{R_w}{R_L+R_w}\right)E_{\text{out}}(n) \qquad (141)$$

This is the same result as obtained in equation 63.

Conclusions

The analysis of the series-connected saturable reactor with external feedback demonstrates that the addition of the rectifier bridge circuit increases the input impedance of the magnetic amplifier, which in turn allows a higher turns-ratio and hence an increased forward gain. Since the gain-bandwidth product is proportional to the voltage gain in the forward transmission, the gain-bandwidth product is thus increased by the addition of the rectifier bridge circuit. From the theory of feedback, it is known that feedback cannot increase the gain-bandwidth product. Therefore, the rectifier bridge circuit does not improve the performance of the magnetic amplifier by causing feedback, but instead by increasing the input impedance. As a result, the present use of the term feedback in magnetic amplifiers is inconsistent with the term feedback used in other fields. In this paper, therefore, the term feedback in magnetic amplifiers has been redefined to make it consistent with already accepted terminology.

It has been shown that all magnetic amplifiers operate upon the same basic principles, and that these principles are demonstrated by the operation of the single-core half-cycle-response magnetic amplifier. By considering the magnetic amplifier as a connection of basic circuits, it is shown that the analysis can be separated into two distinct problems: first, the operation of the basic circuit; and second, the interaction between these circuits resulting from the coupling.

The analysis shows that the time constant in the magnetic amplifier is caused by the inherent time delay in the method of control. This time delay is always present because the flux level in the reactor can only be controlled when the reactor is unsaturated, and a change in the flux level cannot produce a change in the output until the next saturated period. By the use of positive feedback the gain can be increased. However, because of the inherent time delay, the increase in the gain will be matched by a corresponding increase in the time constant.

References

1. Magnetic Amplifiers, **A. G. Milnes.** *Proceedings*, Institution of Electrical Engineers, London, England, vol. 96, pt. I, May 1949, p. 89.

2. Steady-State and Transient Analysis of an Idealized Series-Connected Magnetic Amplifier, **Louis A. Pipes.** *AIEE Transactions*, vol. 70, pt. II, 1951, pp. 2129–35.

3. Transient Response of Saturable Reactors with Resistive Load, **H. F. Storm.** *Ibid.*, pt. I, pp. 95–102.

4. On the Mechanics of Magnetic Amplifier Operation, **Robert A. Ramey.** *Ibid.*, pt. II, pp. 1214–23.

5. The Transductor, D. C. Pre-saturated Reactor, With Special Reference to Transductor Control of Rectifiers, **Uno Lamm.** Esselte Aktiebolag, Stockholm, Sweden, 1943.

6. The Transductor Amplifier, **Ulrik Krabbe** Einar Munksgaard, Copenhagen, Denmark, 1947.

7. Principles of Servomechanisms (book), **G. S. Brown, D. P. Campbell.** John Wiley & Sons, Inc., New York, N. Y., 1948.

8. Transients in Linear Systems (book), **Murray F. Gardner, John L. Barnes.** John Wiley & Sons, Inc., New York, N. Y., 1948.

◆

Discussion

H. H. Woodson (Massachusetts Institute of Technology, Cambridge, Mass.): The author should be complimented on his significant contribution to magnetic amplifier theory. The three major accomplishments reported in this paper should be further emphasized.

1. The correct mathematics (difference equations) to be used with magnetic amplifiers utilizing saturable reactor core materials presently available are shown. The proof of the correct mathematics is contained in the ease of physical interpretation of the results.
2. The use of difference equations in the analysis of magnetic amplifiers puts the performance of all magnetic amplifiers in the same terms. Thus, valid comparison can now be made to allow selection of the best circuit, from the standpoint of dynamic performance, for any application.
3. The paper shows the designer of high-performance servos how the inherent time delay $\epsilon^{-(\pi/\omega)s}$ affects the dynamic characteristics of any magnetic amplifier. This is a distinct aid in design work because a time delay $\epsilon^{-(\pi/\omega)s}$ is considerably more difficult to compensate for than the low signal-frequency approximation $1/[(\pi/\omega)\,s+1]$ in a servo system.

If the author should change his imperfect voltage sources to equivalent imperfect current sources, he would find that some simplification would result in his analysis of the control period. This is especially true for the iterative process indicated by equations 117 to 120. The change to current sources yields considerable simplification when a magnetic amplifier with several control windings is considered. Of course, the end result will be the same as that obtained by the author's methods.

P. R. Johannessen: I agree that the three main objectives should be more clearly stated. Further, the suggestion to change the imperfect voltage sources to imperfect current sources has merit, and may considerably simplify the algebra. By this transformation the circuit can be analyzed by the use of node equations instead of by loop equations. The basic magnetic amplifier (Fig. 1) will now consist of only one node, independent of the number of windings on the reactor. Thus, a single node equation for each period of operation will completely describe the operation of the magnetic amplifier. It is also hoped that the method of analysis presented may be further extended to include a finite slope on the *B-H* curve. I wish to thank Mr. Woodson for his criticism.

The Measurement of Electrostatic Potential Attributable to Net Ion Space Charge in Air

J. S. CARROLL
FELLOW AIEE

S. B. HAMMOND
ASSOCIATE MEMBER AIEE

Synopsis: The electrostatic potential in air resulting from an ion space charge of one predominant sign is difficult to measure with existing potential-measuring instruments. An ideal instrument should not deplete the ion space charge and should be capable of measuring relatively low potentials. An instrument using the null-and-substitution method, and combining an electronic pulse widener and a vacuum-tube voltmeter, is described. This instrument has an application in measuring ion density in air by measuring the electrostatic potential attributable to ion space charge; Poisson's equation can then be solved for ion density.

ALL MEASUREMENT techniques for determining the electrostatic potential in space with respect to ground emphasize the necessity of high internal resistance and high sensitivity in the measuring device. These attributes can be found to a greater or lesser degree in the instruments generally used, such as the gold-leaf electroscope, the electrostatic voltmeter, and the rotary electrostatic voltmeter.

One application of these techniques is the measurement of electrostatic potential in space attributable to a net ion space charge of either negative or positive sign that may exist in the ambient air. By net ion space change is meant the preponderance of one sign over the other, or the difference between the negative and positive space charge. For example, a positive ion net space charge usually exists in a room heated by an unshielded gas or electric heater.[1] Furthermore, considerable interest is being shown in the production and control of net ion densities for physiological reasons.[2]

The requirement of high resistance and sensitivity can be readily understood, for the total number of ions in a space may be very small when compared to the charge flow per second that constitutes 1 microampere. For instance, a room of 1,000 cubic feet (2.83 by 10^7 cubic centimeters) with a net ion density of 500 ions per cubic centimeter has a net total of 1.41 by 10^{10} ions enclosed within its walls. If all of these ions were to move at a uniform rate through one of the walls in 1 second of time, the net current through that wall for 1 second would be

$$I=\frac{dQ}{dt}=0.00228 \text{ microamperes}$$

Thus, any method used to measure the electrostatic potential attributable to the ion space charge must have extreme sensitivity in order to prevent distortion of the field by ion absorption.

The literature on atmospheric electricity studies is an old and extensive one, and has been well summarized in recent years.[3,4] Instruments used in these studies fall roughly into two categories, those that measure the ion density or space charge directly, and those that measure electrostatic potential.

The ion collector, or ion counter, is an example of the first method. Briefly, this instrument collects the ions from the air as it is blown at a known rate between the plates of a charged capacitor. The ion density of the air is then a function of the discharge rate of the capacitor. This method exhibits two distinct shortcomings. First, air must be blown between the capacitor plates at a rate that soon strips an appreciable part of the total ions from the air in an enclosed space or a room. Second, the indicating instrument in the form of an electrometer of considerable sensitivity requires a highly skilled operator to obtain good results.

The electrostatic potential-measuring method has some restrictions because of the particular instrument used. Electrometers possessing the necessary voltage sensitivity and high input resistance are, by nature, extremely delicate instruments requiring laboratory technique and skill of a high order. Their movements are slow-acting, and generally they are not rugged enough to be portable or readily set up for measurements. However, they can be made to give good results.

With these restrictions in mind, the instrument to be described represents an effort to apply new techniques to an old problem, with the hope that further improvements and advancements will follow.

Method of Measurement

Any high-impedance voltage-measuring instrument, such as the gold-leaf electroscope or the electrostatic voltmeter, can be adapted to measure electrostatic potentials in space. A suitable probe is required, the purpose of which is to raise or lower the potential with respect to ground of one terminal of the instrument to coincide with the actual potential in space, measured with respect to ground. A usable system employs a gold-leaf electroscope and radioactive probe, as shown in Fig. 1. The null method may be employed with this system, as the diagram indicates, or the electroscope may be made to read directly in volts by grounding the case and calibrating the electroscope with an adjustable direct-voltage source. The latter method tends to distort the potential field of the space charge by placing the grounded case close to the probe. The null method corrects this distortion by maintaining the electroscope case at a potential equal to the potential in space attributable to the space charge. When the battery voltage is adjusted so that the electroscope reads zero, the potential of the probe can be read on the voltmeter.

The null method has the advantage of restricting the field distortion of the grounded case. A further advantage exists in the reduction of ion depletion in the space charge field by leakage. If the electroscope case is at ground potential, and the probe is at a potential with respect to ground other than zero, a potential gradient exists along the polystyrene insulation to the case. This resistance is very high and has been measured at 6× 10^{15} ohms for one specially built electro-

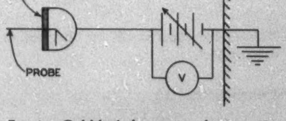


Fig. 1. Gold-leaf electroscope for measuring electrostatic potential in space

---

Paper 54-511, recommended by the AIEE Basic Sciences Committee and approved by the AIEE Committee on Technical Operations for presentation at the AIEE Fall General Meeting, Chicago, Ill., October 11–15, 1954. Manuscript submitted October 20, 1953, made available for printing August 24, 1954.

J. S. CARROLL is with Stanford University, Stanford, Calif. S. B. HAMMOND is with the University of Utah, Salt Lake City, Utah, and was with Stanford University at the time the work was done on the paper.

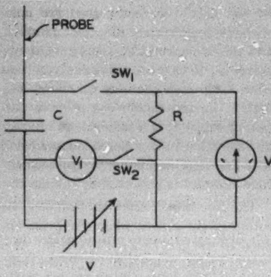


Fig. 2. Elementary circuit

scope, but it may yet conduct appreciable current. By the null method, however, no potential gradient exists across the insulation, and hence there is no leakage of charge from the probe to ground.

The gold-leaf electroscope must be read from a distance through a telescope in order that the observer does not distort the electric field of the space charge which the electroscope is measuring. An additional disadvantage of the gold-leaf electroscope is readily apparent to anyone who has attempted to obtain quantitative results with one; great care and delicate technique are necessary. However, the leakage resistance can be made higher than that of a commercial electrostatic voltmeter by a factor of over 100. Furthermore, the speed of response is greater. The gold-leaf electroscope can be made to give good results under conditions of fairly high potentials.

Radioactive Probe

The purpose of the probe is to establish and maintain one terminal of the electrostatic potential-measuring device at the potential of the surrounding space. For some applications the historic water-dropper or a small flame may be used. However, where the electric field is produced mainly by the ion space charge, and where the space charge may be either positive or negative in sign, the radioactive probe is more desirable. Polonium is often used as the radioactive material. Essentially an alpha emitter, polonium need only be used in very small quantity. The charged alpha particles emitted into the space around the probe have a negligible effect on the existing ion space charge field. Yet each alpha particle produces about 1.6×10^5 ion-pairs in the air surrounding the probe. The range of alpha particles is about 3.84 centimeters in air at atmospheric pressure.[5] The ion pairs have no over-all effect on the space-charge field, but provide a sheath of high conductivity in the vicinity of the probe. If a potential difference exists between the probe and the surrounding space, a potential gradient will exist at the surface of the probe. Charge will flow in this high-conductivity region within the ion sheath between the space and the probe, maintaining the probe at the potential of the space.

If there is a leakage of charge from the probe to ground, it must be supplied continuously by the space charge field under study. An undesirable decrease in the initial field strength naturally results. Furthermore, the constant leakage of charge maintains the probe potential somewhat below the potential of the surrounding space by the voltage drop necessary to supply the leakage ion current. This results in an additional error in measuring the potential. For these two reasons, the probe should have to supply zero current to the potential-measuring instrument, if the probe is to be maintained at the same potential which the space would have without the probe. But since leakage current always exists, it should be kept at a minimum.

Vacuum-Tube Electrometer

Vacuum-tube electrometers generally possess a limit of sensitivity which is of the order of 10^{-12} amperes full scale and 10^{11} ohms. Compared to the conventional vacuum-tube voltmeter, the electrometer is extremely sensitive, but for the type of measurements under discussion, it is hardly satisfactory. Yet an electronic instrument would have some advantages over the gold-leaf electroscope, which is unreliable for small potentials, and is generally temperamental.

Elementary Circuit

The instrument to be described combines an energy storage system with the null-and-substitution method. The basic principle of operation can be seen in Fig. 2. Assume that the probe has had time to acquire the same potential that the point in space would have if the probe were not present. Capacitor C is charged to the potential of the probe. Substitution voltage V is adjustable so that if the voltage V is equal to the potential on C, no current flows through R when switch 1 is closed. This is the null condition, and the voltage V may be read on voltmeter V_1 by closing switch 2. If, however, the battery voltage is greater or less than the potential of the charge on C, an exponential current pulse of negative or positive sign respectively will flow through R. The voltmeter V_2 has a zero center reading, and will show the sign of the voltage pulse across R. Thus it is possible to determine whether battery voltage V is greater or less than the voltage on C by the negative or positive swing of the voltmeter V_2. For this part of the discussion it has been assumed that the time constant of the exponential voltage pulse is sufficient that the voltmeter V_2 will have time to show a displacement for each pulse. Successive adjustments of battery voltage V can be made, each time closing switch 1 for an instant to determine whether V is greater or less than the value necessary to balance the capacitor voltage. In this way a null condition can be approached and reached, and the potential of the probe can be read on voltmeter V_1 through the intermediary of the substitution voltage V.

The time constant RC of Fig. 2 will ordinarily be of the order of 100 micro-

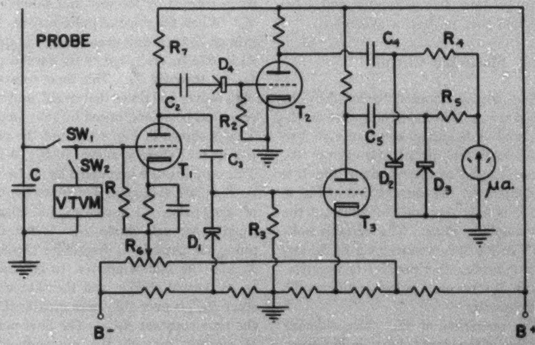


Fig. 3. Basic electronic circuit

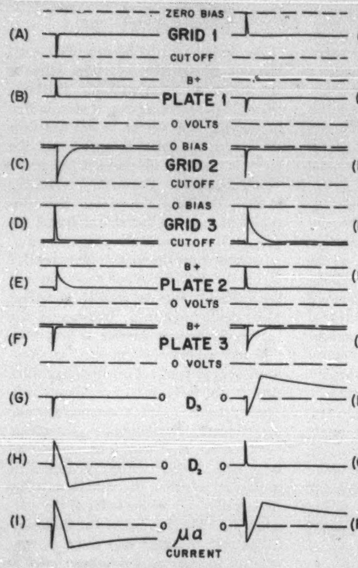


Fig. 4. Voltage wave forms during operation. Left—probe voltage less than substitution voltage. Right—Probe voltage greater than substitution voltage

seconds, which is too short a time for voltmeter V_2 to show an indication of the pulse. The value of C, which consists entirely of the stray capacitance of the probe and circuit wiring, should be kept small so that the probe can charge in a reasonable time. R can be made as large as possible, but in this simplified circuit the resistance of voltmeter V_2 must be considered, and it will be the limiting value of R. The value of C can be about 5 to 10 micromicrofarads, and R can be about 10 to 20 megohms. With a time constant of 10^{-4} seconds, it becomes necessary to lengthen the duration of the pulse so that the voltmeter will have sufficient time to show a deflection.

Basic Electronic Circuit

The method adopted to lengthen the time constant is included in a pulse amplifier, and it lengthens the pulse in two stages, each time by a factor of over 100. The operation of the circuit of Fig. 3 is fundamentally the same as that of Fig. 2, but R is now the grid resistance of the first amplifier stage. The variable voltage V of Fig. 2 is produced across R_6, and may be made either positive or negative. Triode T_1 serves as a pulse amplifier and a pulse inverter.

The operation of the pulse-widening function of the circuit shown in Fig. 3 can best be described by following a pulse through the entire circuit. Assume first that the potential across C is less positive than that from the grid of T_1 to ground. When Sw_1, a push-button switch, is closed a negative pulse appears on the grid of T_1 because of the charging current flowing through R to C. T_1 amplifies and inverts this pulse, which is coupled to the grids of T_2 and T_3 by means of C_2 and C_3 respectively. T_3 amplifies the pulse and passes it on with the same time constant, but the operation of T_2 lengthens the time constant of this pulse. The grid of T_2 is made positive during the pulse, and C_2 is able to increase its charge rapidly through the diode formed by the grid and cathode of T_2. When the original pulse is over, the grid of T_2 is carried negative by C_2, until C_2 is able to discharge to its normal potential through R_2. The time constant R_2C_2 is over 100 times that of RC, and the pulse has been lengthened by that factor. This is shown in Fig. 4(C), and the output of that stage is shown in Fig. 4(E). This output pulse is coupled by R_4 by C_4. As the plate voltage of T_1 follows the top of the pulse, C_4 increases its charge rapidly through diode D_2. When the pulse has passed, C_4 discharges through R_4 and the microammeter, to get back to its normal charge, and the time constant R_4C_4 is over 100 times greater than the time constant R_2C_2. The final pulse of current through the microammeter caused by the pulse through T_2 is shown in Fig. 4(H). T_3 has passed the pulse without lengthening its time constant, and this is shown in Fig. 4(G). The two currents add to produce the final resultant current through the microammeter, Fig. 4(I). The area under the positive portion of Fig. 4(I) is insufficient to produce an indication on the microammeter, and only the negative pulse of several seconds' time constant is indicated by the meter.

The foregoing discussion describes the action of the circuit when the pulse applied to the grid of T_1 is negative. If a positive pulse is applied, i.e., if the voltage from grid to ground is less positive than the capacitor voltage, the path of the pulse through T_3 lengthens the pulse by the diode action of D_1 and D_3. T_3 lengthens the pulse by the diode action of D_1 and D_3. Triode T_3 is ordinarily biased to cutoff. A negative pulse from the plate of T_1 drives the grid of T_3 more negative in potential and allows C_3 to discharge rapidly through D_1. When the pulse from T_1 has passed, C_3 pulls up the grid voltage of T_3 in voltage, and C_3 discharges through R_3 with a time constant C_3R_3. Thus the time constant has been increased by a factor of over 100. The curves on the right side of Fig. 4 indicate these operations. Similar action by C_5, R_5, and D_3 lengthen the pulse by another factor of over 100, and the resulting pulse is indicated as a current flowing through the microammeter.

In the preceding discussion it must be remembered that resistance R_3, R_4, and R_5 have in parallel with them the back resistances of the diodes D_1, D_2, and D_3. In calculating the time constants, these parallel resistances must be taken into account. Also, considerable clipping action may result in the triodes when pulses of large amplitude are applied. This has no effect on the practical operation of the circuit.

The vacuum-tube voltmeter shown in Fig. 3 for measuring the substitution voltage must have a very high input resistance in order to measure the substitution voltage across R_6 and the grid resistance R. A commercial instrument of 100-megohm input resistance is satisfactory, provided R is not greater than about 20 megohms. The instrument should have, in addition, voltage scales sufficient to cover the range of potentials which R_6 is capable of producing.

Sensitivity

There are two means by which this electrometer may deplete the ion space charge in the space being studied. First, there is a constant leakage in the insulation from the probe circuit to ground. This may be minimized by the use of polystyrene for any insulators needed for supporting the probe and lead-in wire, and for insulation for switch 1, which must be specially constructed for long leakage paths. Second, each time capacitor C is discharged through R, a quantity of charge must be extracted from the space charge to recharge C. By keeping C small and allowing sufficient time for it to charge after each discharge, the absorption effect on the space charge field may be minimized. When the voltage across R_s is close to balance with the voltage of capacitor C, only a small part of the charge on C is lost when switch 1 is closed. By systematically adjusting R each time an attempt at balance is made, a balance can be approached with a minimum number of attempts.

One advantage of this instrument is its ability to read small voltages without relying on the calibration of the input resistance. The authors have had considerable success with a range of zero to plus or minus 20 volts. By including an additional stage of pulse amplification, this instrument can be operated for full-scale deflection of 1 volt or less.

Applications

While the measurement of electrostatic potentials in space may be of interest, a further application of the instrument is to be found in converting this potential reading, if made in an enclosed space, to net ion density. This can be accomplished by solving Poisson's equation for the particular boundary values. The solution of Poisson's equation is very easy for simple geometries, such as infinite parallel planes, a sphere, or a cylinder of infinite length. Approximation of existing boundaries to these simple ones can often be made, but for other configurations the solution usually becomes very much involved, and often it is necessary to resort to numerical approximation methods.

References

1. CLIMATE CONTROL THROUGH IONIZATION, **Thomas L. Martin.** *Journal*, Franklin Institute, Philadelphia, Pa., vol. 254, Oct. 1952, pp. 267–80.

2. CONTROL OF AIR ION DENSITY IN ROOMS, **H. H. Skilling, John C. Beckett.** *Ibid.*, vol. 256, Nov. 1953, pp. 423–34.

3. ATMOSPHERIC ELECTRICITY (book), **J. Alan Chalmers.** Clarendon Press, Oxford, England, 1949.

4. TERRESTRIAL MAGNETISM AND ELECTRICITY (book), edited by **J. A. Fleming.** Dover Publications, New York, N. Y., 1949.

5. INTRODUCTION TO ATOMIC PHYSICS (book), **Henry Semat.** Farrar and Rinehart, Inc., New York, N. Y., 1939, p. 269.

◆

Discussion

T. W. Liao (University of Utah, Salt Lake City, Utah): The authors have presented a practical method capable of measuring relatively low potentials in air resulting from an ion space charge of one predominant sign.

By means of the null-and-substitution method they have been able to overcome the usual high resistance and sensitivity requirements of the measuring devices of this kind.

While this measuring technique might have many applications in the determination of ion densities in air, it appears to have possible applications in the measurements of space charges in insulating liquids such as transformer oil. In the latter case, similar limitations, as discussed in the paper, are imposed.

The radioactive probe technique might also be of advantage to the measurements in liquids in that it reduces the potential difference between the probe and the surrounding liquids. Because of the higher density of liquids relative to air, the high-conductivity sheath would be thinner. This should result in a better resolution in the measurement of the space charge distribution than that which can be obtained in a gaseous medium.